

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 09:54:26 ; Search time 86.5 Seconds
(without alignments)
9941.972 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 4896

Sequence: 1 aatctttattttatgatg.....aggcttttttctcctaataacc 2709

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2.1/USPTO spo01/US09774490/runat 03082003 095415 25314/app query.fasta_1.2887
-DB=J_Geneseq_19Jun03 -QWMT=fastan -SUFFIX=reg -MINMATCH=0.1 -LOOPFCU=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=BLOSUM62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09774490 @CEN 1.114 @runat 03082003 095415 25314 -NCPU=6 -ICPU=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THRRADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A Geneseq 19Jun03.*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	4201	85.8	771	16	AA71380	Human semaphorin I
2	4201	85.8	771	22	AAG62726	Amino acid sequenc
3	4201	85.8	771	23	ABG96413	Human ovarian canc
4	4194	85.7	796	19	ABY21264	Human semaphorin I
5	2562	52.3	477	16	AA741175	Human collapsin.
6	2245.5	45.9	777	20	AA727127	Human brain tissue
7	2245.5	45.9	777	21	AA799427	Human PRO1491 (UNQ
8	2245.5	45.9	777	22	AAU29197	Human PRO polypept
9	2245.5	45.9	777	22	AAB66176	Protein of the inv
10	2245.5	45.9	777	24	ABU71285	Human PRO1491 prot
11	2245.5	45.9	777	24	ABU65742	Human secreted/tra
12	2245.5	45.9	777	24	ABU66075	Novel human secret
13	2245.5	45.9	777	24	ABU67579	Human secreted/tra
14	2245.5	45.9	777	24	ABU65437	Human PRO polypept
15	2245.5	45.9	777	24	ABU58573	Human PRO polypept
16	2245.5	45.9	777	24	ABU56109	Human secreted/tra
17	2245.5	45.9	777	24	ABU57104	Human PRO polypept
18	2245.5	45.9	777	24	ABU10683	Human secreted/tra
19	2209.5	45.1	807	22	AAG62729	Amino acid sequenc
20	2118	43.3	749	22	AAG62727	Amino acid sequenc
21	2044	41.7	775	24	ABR47588	Breast cancer asso
22	2030	41.5	775	19	AAW63748	Human semaphorin.
23	2024	41.3	775	20	AA743090	Mouse semaphorin H
24	1942	39.7	777	20	AA743091	Mouse semaphorin H
25	1942	39.7	777	22	AAG62731	Amino acid sequenc
26	1937	39.6	785	22	AAG62730	Amino acid sequenc
27	1849.5	37.8	751	20	AAW30617	Human semaphorin E
28	1849.5	37.8	751	21	AAW28379	Clone BR533.4. Ho
29	1849.5	37.8	751	23	ABP68623	Human pancreatic c
30	1847.5	37.7	751	22	AAG62728	Amino acid sequenc
31	1834	37.5	875	22	AAG65619	Novel human protei
32	1830	37.4	782	21	AAW23609	Human secreted pro
33	1830	37.4	782	22	AAG65620	Novel human protei
34	1826.5	37.3	785	21	AAW23636	Human secreted pro
35	1751.5	35.8	779	22	ABW84219	Amino acid sequenc
36	1751.5	35.8	779	23	AAW78481	Human ZSMF-16. Ho
37	1010	20.6	861	18	AAW17658	Mouse CD100 antige
38	1003	20.5	861	19	AAW58540	Human semaphorin.
39	1003	20.5	861	22	AAW1035	Murine CD100 amino
40	1003	20.5	861	22	AAW1251	Mouse CD100 protei
41	981.5	20.0	862	18	AAW17657	Human CD100 antige
42	981.5	20.0	862	22	AAW1036	Human CD100 amino
43	981.5	20.0	862	22	AAW1252	Human CD100 protei
44	966	19.7	834	22	AAW61238	Murine M-Sema-F pr
45	957.5	19.6	833	22	AAW03640	Human extracellular

ALIGNMENTS

RESULT 1

AA71380
ID AA71380 standard; Protein; 771 AA.

AC AA71380;

XX 25-MAR-2003 (updated)

DT 21-NOV-1995 (first entry)

XX Human semaphorin III protein.

XX Semaphorin; grasshopper; human; vaccinia virus; Drosophila; Tribolium;

XX variola major virus; smallpox; semaphorin receptor binding activity;

XX modulation; nerve cell growth; immune response; viral pathogenesis;

XX neurological disease; neuro-regeneration; oncological infection.

XX Homo sapiens.

OS XX

PN W09507706-AL.
 XX
 PD 23-MAR-1995.
 XX
 PF 13-SEP-1994; 94WO-US10151.
 XX
 PR 13-SEP-1993; 93US-0121713.
 XX
 PR (REGC) UNIV CALIFORNIA.
 PA
 XX Bentley DR, Goodman CS, Kolodkin AL, Matthes D;
 PI O'Connor T;
 XX WPI; 1995-131177/17.
 DR N-PSDB; AAQ87442.
 DR
 XX
 PT New class of semaphorin peptide(s) and polypeptide(s) - are
 PT potent modulators of nerve cell growth and regeneration
 XX
 PS Example 2; Page 60-63; 101pp; English.
 XX
 CC The sequence of the human semaphorin III protein. The proteins
 CC encoded by the grasshopper semaphorin I (AAQ87441), human semaphorin
 CC III, vaccinia virus semaphorin IV (AAQ87443), Drosophila semaphorin I and
 CC II (AAQ87444-5), Tribolium semaphorin I (AAQ87446) or variola major
 CC (smallpox) virus semaphorin IV (AAQ87447) genes were used to generate a
 CC series of peptides (AA070370-R70418), which retain semaphorin receptor
 CC binding activity. The semaphorin derived or semaphorin receptor derived
 CC peptides are potent modulators of nerve cell growth, immune
 CC responsiveness and viral pathogenesis. They can be used in diagnosis and
 CC treatment of neurological disease and neuro-regeneration, immune
 CC modulation and diagnosis and treatment of viral and oncological infection
 CC and diseases.
 CC (Updated on 25-MAR-2003 to correct PN field.)
 XX
 SQ Sequence 771 AA;

Alignment Scores:
 Pred. No.: 0 Length: 771
 Score: 4201.00 Matches: 771
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 85.80% Indels: 0
 DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x AAR71380 (1-771)

QY 200 ATGGGCTGTTAACTAGGATTGTCTGCTTTTCTGGGAGTATTACTTACAGCAAGAGCA 259
 |||
 Db 1 MetGlyTyrLeuThrArgileValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
 |||
 QY 260 AACTATCAGAATGGGAGAACAAATGTGCCAGCGTGAATATCTTACAAAGAAATGTG 319
 |||
 Db 21 AsnTyrGlnAsnGlyLysAsnValProArgLeuLeuLeuSerTyrLysGluMetLeu 40
 |||
 QY 320 GAATCCAAACATGTGATCACTTTCAATGCTTGGCCCAACAGCTCCAGTATCATCACTTC 379
 |||
 Db 41 GluSerAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
 |||
 QY 380 CTTTGTGGATGAGGACGGAGTAGGCTGTATGTGTGGAGCAAGAGATCATATTTTCATTC 439
 |||
 Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
 |||
 QY 440 GACCTGGTTAAATATCAGGATTTTCAAAGATTGTGGCCAGTATCTTACACCAAGA 499
 |||
 Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
 |||
 QY 500 GATGAATCAAGTGGCTCGAAAGACATCTCGAAAGATGTGCTAAATTTTCATCAAGTGA 559
 |||
 Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPhelleyVal 120
 |||
 QY 560 CTTAAGGCATATAATCAGACTCACTTGTCGCTGTGGAAACGGGGGCTTTTCATCCAATT 619
 |||

Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
 |||
 QY 620 TGCACCTACATTGAAATTGGACATCATCTCAGGACATATATTTTAACTCGAGACTCA 679
 |||
 Db 141 CysThrTyrIleGluIleGlyHisProGluAspAsnIlePheLysLeuGluAsnSer 160
 |||
 QY 680 CATTTTGAACCGCGGTGGGAAGAGTCCATATGACCTTAAGCTGCTCAGAGCATCCCTT 739
 |||
 Db 161 HisPheGluAsnGlyArgGlySerProTyrAspProLysLeuLeuThrAlaSerLeu 180
 |||
 QY 740 TTAATAGATGAGAAATTATATCTCTGGAACTCGAGCTGATTTTATGGGCGGAGACTTGT 799
 |||
 Db 181 LeuIleAspGlyLeuLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
 |||
 QY 800 ATCTTCCGAACCTCTGGCACACCCATCAGGACAGGACAGCATGATTCAGAGTGG 859
 |||
 Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
 |||
 QY 860 CTCATATGATCCAAAGTTTCATTAGTGCCACCTCATCTCAGAGAGTGACAATCCTCAAGAT 919
 |||
 Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuLeuSerGluSerAspAsnProGluAsp 240
 |||
 QY 920 GACAAAGTATATCTTTTCTCCGTCAAAATCAATAGATGAGAACACTCTCGAAAGACT 979
 |||
 Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
 |||
 QY 980 ACTACGCTAGAAATAGTGCAGATATGCAGATACATCTTGGAGGCGACAGAGTCTGGTG 1039
 |||
 Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
 |||
 QY 1040 AATAAATGGACAACATCTCTCAAAAGCTCGTCTGATTTTCTCAGTGCCAGAGTCCAAATGCG 1099
 |||
 Db 281 AsnLysTrpThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
 |||
 QY 1100 ATTCACTCATCTTTTGTGATGAACACTGCAGATGTATCTCTAATGAATTTTAAAGATCTCTAAA 1159
 |||
 Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
 |||
 QY 1160 AATCCAGTTGTATATGGAGTGTATTCAGACTTCCAGTAACTATTTCAAGGGATCAGCCGTG 1219
 |||
 Db 321 AsnProValValTyrGlyValPheThrSerSerAsnIlePheLysGlySerAlaVal 340
 |||
 QY 1220 TGTATGTATAGCATGAGTGTGAGAGGGGTGTTCTTGTGCTCATATGCCACAGGGAT 1279
 |||
 Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
 |||
 QY 1280 GGACCCCAACTATCAATGGTGCCTTATCAAGAGAGTCCCTATCCACGCCAGGAACT 1339
 |||
 Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
 |||
 QY 1340 TGTCCCGACCAACATTTGGTGGTTTGTACTCTCAAGGACCTTCTCTGATGATGTTATA 1399
 |||
 Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
 |||
 QY 1400 ACCTTTGCAAGAAGTCAATCCAGCCATGTACAATCCAGTGTCTTCTATGAAACAATCGCCA 1459
 |||
 Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
 |||
 QY 1460 ATAGTGATCAAAACGGATGTAATATCAATTTACAAATTTGCTGTAGACCGAGTGTGAT 1519
 |||
 Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
 |||
 QY 1520 GCAGAGATGGACAGTATGATGTTATCTTTATCGGACAGATGTTGGACCGTCTCTTAAA 1579
 |||
 Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
 |||
 QY 1580 GTAGTTTCAATTCCTCAGGAGACTTGTGTATGATTTAGAGAGGTTCTCTCGAAGAAATG 1639
 |||
 Db 461 ValValSerIleProLysGluThrTyrTrpTyrAspLeuGluValLeuLeuGluGluMet 480
 |||
 QY 1640 ACAGTTTTCGGGAACCGACTGCTGCTATTTTCAGCAATGAGAGCTTTCCATCAAGCAGCAAA 1699
 |||
 Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500

QY 1700 CTATATATTGGTTCAACGGCTGGGTTGCCAGCTCCCTTTTACACCGGTGTGATATTAC 1759
Db |||||
501 LeuTyriLeGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGCGTGTGCTAGTGGTTCCTCGCCGAGACCTTACTGTGCTGGATGCTTCT 1819
Db |||||
521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCCACTGCAAGAGACGCACAAAGACGACGAAGATAAAGAAAT 1879
Db |||||
541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTCTGTTACAGCTTACACCATGATAATCATCATGCGCCACAGCCCT 1939
Db |||||
561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580
QY 1940 GAAGAGAGATCATCTATGCTGTAGAGATAGTACACATTTTTCGAATGCAAGTCCGAG 1999
Db |||||
581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGACCGCTGCTCTATTGGCAATTCAGAGCGAAATGAAGAGCGAAAGAGAG 2059
Db |||||
601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGlu 620
QY 2060 ATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCCTTCTCTAGTGTCTACAA 2119
Db |||||
621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAGAGTTCAGCAATACCTCTGCCATCGGTGGACATGGGTTTCATACAACTCTT 2179
Db |||||
641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAAACCTCGAAGTCTATTCACACAGAGCAATTTGGAAGAACTTCTTCATAAGAT 2239
Db |||||
661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLysAsp 680
QY 2240 GATGATGAGATGGCTTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAGAG 2299
Db |||||
681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 CTCTGTCACAGAGTCTCATGCTGCTCATCAACACCCCACTCTCAACACAGTGTAG 2359
Db |||||
701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAGTTTGGAAAGGACCGAAACAAACCTCGCAAGGCCAGGACATACC 2419
Db |||||
721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGAAACAGTAAACAAATGGAAGCACTTACAGAAATAAAGAGGTAGAAACAGGAGG 2479
Db |||||
741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGlnAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCCAGCAATTGAGAGGCGCCACAGGAGTGTCT 2512
Db |||||
761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 2
ID AAG62726
XX
AC AAG62726;
XX
DT 17-SEP-2001 (first entry)
XX
DE Amino acid sequence of human semaphorin Sema3A.
XX
KW Hanatoxin; tarantula; toxin; voltage-gated potassium channel;
KW voltage-gated calcium channel; hanatoxin-like sequence; HTLS;
KW semaphorin; dorsal root ganglion repulsion; growth cone collapse.
XX
OS Homo sapiens.
XX

PN WO200138491-A2.

XX 31-MAY-2001.

XX 07-NOV-2000; 2000WO-US41943.

XX 08-NOV-1999; 99US-0164056.

XX (GBHO) GEN HOSPITAL CORP.

XX Behar O, Woolf CJ;

XX WPI; 2001-451494/48.

Polypeptide sequences that encompass the hanatoxin-like sequences of semaphorins, useful as drugs to treat any condition or disease that is characterized by abnormal calcium channel function -

Claim 6; Page 11; 29pp; English.

The present sequence represents a semaphorin. The specification describes hanatoxin-like sequences (HTLS) found in the semaphorin domain of mammalian secreted semaphorins. Hanatoxin is a tarantula toxin that selectively blocks some voltage-gated potassium and calcium channels. The HTLS is responsible for the dorsal root ganglion repulsion and growth cone collapse activities associated with semaphorins. Polypeptides containing HTLS can be used to modulate the activity of calcium channels. The peptides can also be used as an antigen to generate antibodies that can then be used to modulate the activity of calcium channels by inactivating naturally occurring channel ligands. The peptides or antibodies can be used as drugs to treat any condition or disease that is characterized by abnormal calcium channel function.

SQ Sequence 771 AA;

Alignment Scores:

Pred. No.:	0	Length:	771
Score:	4201.00	Matches:	771
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	85.80%	Indels:	0
DB:	22	Gaps:	0

US-09-774-490-1 (1-2709) x AAG62726 (1-771)

QY 200 ATGGGCTGGTTAACTAGGATTGTCTGCTTTTCTGGGAGTATTACTTACAGCAGAGCA 259

Db ||||| 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20

QY 260 AACTATCAGAAATGGGAAGAACAAATGTGCAAGGCTGAAATTTATCTCAAAAGAAATGTTG 319

Db ||||| 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40

QY 320 GAATCCAAACAAATGTGATCATTTCATCAATGGCTGGCCAAACAGCTCCAGTTATCATCCCTC 379

Db ||||| 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyrHisThrPhe 60

QY 380 CTTTGGATGAGGACGAGTGTATCTTCTGGACGAAAGGATCATATTTTCATTC 439

Db ||||| 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80

QY 440 GACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGGCCAGTATCTTACACCAAGAGA 499

Db ||||| 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100

QY 500 GATCAATGCAAGTGGGTGGAAAGACATCTCTGAAAGAAATGTGCTAATTTTCATCAAGTA 559

Db ||||| 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120

QY 560 CTTAAGGCATATATCAGACTCAGCTTCTAGCCTGTGGAACGGGGCTTTTCATCCCAATT 619

Db ||||| 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140

620 TGCACCTACATTGAATGGACATCATCTCTGAGGACAAATATTTTAAGCTGGAGAACTCA 679
141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePhePheLeuGluAsnSer 160
680 CATTTTGAACCGCGTGGAGAGTCCATATGACCCCTTAAGCTCTCTGACAGCATCCCTT 739
161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
740 TTAATAGATGAGAAATTAATCTCTGGAAGTCTGAGCTGATTTTATGGCGGAGACTTTCCT 799
181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
800 ATCTTCCGAACTCTGGGACACACCAATCAGACAGACAGCATGATTCCAGGTGG 859
201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTyr 220
860 CTCAATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGCAATCCTCAAGAT 919
221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
920 GACAAAGTATATCTTTCTTCCTGCAAAATGCAATAGATGGAGAACCTCTGGAAAGCT 979
241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
980 ACTCAGCTAGAAATAGGTGAGATATGCAAGATGACTTTGGAGGCGCAGAGTCTGGTG 1039
261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
1040 AATAAATGGACAACTTCTCAAGCTCGTCTGATTGCTCAGTCCAGGTCCTCAATGCG 1099
281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
1100 ATTGACACTATTTGATGAAGTCTGAGATGATGATTCCTTAATGAATTTAAGATCCTAAA 1159
301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
1160 AATCAGTTGTATATGGAGTGTTCAGACTTCAGTAACTTTCAAGGATCAGCCGTG 1219
321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
1220 TGTATGTATAGATGATGATGAGAGGGTGTTCCTTGGTCCATATGCCACAGGAT 1279
341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
1280 GCACCAACTATCAATGGTGCTCTATCAAGAGAGTCCCTATCCAGCGGCAGAACT 1339
361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
1340 TGTCACGCAAAACATTTGGTGGTTTGTACTCTACAAGGACCTTCCTGATGCTTATA 1399
381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
1400 ACCTTTGCAAGAGTCAATCCAGCCATGTACATCCAGTGTTCCTATGAACAATCGCCCA 1459
401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
1460 ATAGTATCAAAACCGATGTAATTAATCAATTTACAAATTTGCTGAGACCGAGTGGAT 1519
421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
1520 GCAGAGATGGACAGTATGATGTTTGTATGTTTATCGGAACAGATGTTGGGACCCCTCTAAA 1579
441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTTAGAAGAGTTCCTGCGAAGAAATG 1639
461 ValValSerIleProLysGluThrTrpTyrAspLeuGluGluValLeuLeuGluMet 480
1640 ACAGTTTTCGGAACCGACTGCTATTTCCAGAAATGGAGCTTTCCTCACTAAGCAGCAACA 1699
481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
1700 CTATATATTGGTTCACGCGCTGGGGTTGGCCAGCTCCCTTTTACACCGGTGTGATTTTAC 1759

501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
1760 GGGAAACGCTGCTGCTAGTGTTCCTGCCCGACACCTTACTGCTGCTGGATGGTTCCT 1819
521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
1820 GCATGTTCTCGCTATTTTCCACTGCAAGAGAGCGCAACAGCACCAAGATATAAGAAAT 1879
541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn 560
1880 GGAGACCACTGACTCTACTGTTTACAGCTTACCATGATATCAATCCATGGCCACAGCCCT 1939
561 GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisGlyHisSerPro 580
1940 GAAGAGAGATCATCTATGTTGTAGAGAAATAGTAGACACATTTTGGATGAGCCCGAAG 1999
581 GluGluArgIleIleTyrGlyValGluAsnSerThrPheLeuGluCysSerProLys 600
2000 TCGCAGAGAGCGCTGCTTATTGGCAATTCAGAGGCGCAATGAAGACGCAAAAGACAG 2059
601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgGlnGluArgLysGluGlu 620
2060 ATCAGAGTGGATGATCATATCATCAGACAGATCAAGCCCTTCTGCTACGTAGTCTACA 2119
621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
2120 CAGAAGATTCCAGCAATTTACTCTGCTGATGCGGTGGAACATGGGTTCATACAACTCTT 2179
641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
2180 CTTAAGGTAACTCCGGAAGTCATTGACACAGAGCATTTGGAAGAACTCTTTCATAAAGAT 2239
661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLysAsp 680
2240 GATGATGGAGATGCTCTTAAGACCAAAAGAAATGTCATATGACATGACACCTTAGCCAGAAG 2299
681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
2300 GTCTGGTACAGACTTCTATGCTGCTCATCATCACCCCAATCTCAACACATGATGATGAG 2359
701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
2360 TTCTGTGAACAAGTTTCGAAAAGGACCGAAACAAACCTCGGCAAGCCAGGACATACC 2419
721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgArgGlnArgProGlyHisThr 740
2420 CCAGGGAACAGTAAACAAATGGAAGCCTTACAGAAATAAGAAAGGTAGAAAACAGGAGG 2479
741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
2480 ACCACGATTTGAGAGGGGACCCAGGAGTGTCT 2512
761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 3
ABG96413
ID ABG96413 standard; Protein; 771 AA.
AC ABG96413;
XX
DT 11-DEC-2002 (first entry)
XX
DE Human ovarian cancer marker M473.
XX
Human; ovarian cancer; marker; cancer; familial history; brain disorder;
central nervous system disorder; bacterial meningitis; viral meningitis;
Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
brain herniation; inflammation; encephalitis; testicular disorder;
nontuberculous granulomatous orchitis; connective tissue disorder;
heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
histological type; carcinogenic; ovarian cancer marker.

OS Homo sapiens.
 XX WO200271928-A2.
 XX 19-SEP-2002.
 XX 14-MAR-2002; 2002WO-US07826.
 XX 14-MAR-2001; 2001US-276025P.
 PR 14-MAR-2001; 2001US-276026P.
 PR 10-AUG-2001; 2001US-311732P.
 PR 19-SEP-2001; 2001US-323580P.
 PR 26-SEP-2001; 2001US-324967P.
 PR 26-SEP-2001; 2001US-325102P.
 PR 26-SEP-2001; 2001US-325149P.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;
 PI Meyers RE, Morrisey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;
 PI Bast RC, Lu K, Schmandt RE, Zhao X, Glatt K;
 XX WPI; 2002-723277/78.
 DR N-PSDB; ABS76512.
 XX Assessing whether a patient is afflicted with ovarian cancer, useful in
 PT assessing the stage or progression of the disease, comprises comparing
 PT the expression level of a cancer marker in a sample from a patient and
 PT from a non cancer patient -
 XX Disclosure; Page 401-402; 481pp; English.
 XX The present invention relates to a new method for assessing whether a
 CC patient is afflicted with ovarian cancer. The method involves comparing
 CC the expression level of a marker in a patient sample and the normal level
 CC of expression of the marker in a control non-ovarian cancer sample, where
 CC the marker is selected from 363 cancer markers described in the
 CC specification. The method of the invention is useful in diagnosing or
 CC characterizing cancer, in detecting the presence of cancer as early as
 CC possible, and the recurrence of ovarian cancer. The method may also be of
 CC particular use with patients having an enhanced risk of developing
 CC ovarian cancer (e.g. patients having a familial history of ovarian
 CC cancer). The cancer markers may be used in the management and treatment
 CC of e.g. brain and central nervous system disorders (e.g. bacterial and
 CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain
 CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
 CC inflammations (e.g. bacterial or viral meningitis or encephalitis),
 CC testicular disorders (e.g. nontuberculous granulomatous orchitis),
 CC connective tissue disorders, or heart disorders (e.g. ischaemic heart
 CC disease or atherosclerosis). The compositions and methods may also be
 CC used in assessing the histological type of neoplasm associated with
 CC ovarian cancer, monitoring the progression of ovarian cancer,
 CC determining whether ovarian cancer has metastasized or is likely to
 CC metastasize, selecting a composition for inhibiting ovarian cancer,
 CC assessing the ovarian carcinogenic potential of a compound, or
 CC inhibiting ovarian cancer or at risk of developing ovarian cancer. The
 CC present amino acid sequence represents one of the ovarian cancer markers
 CC described in the invention.
 XX Sequence 771 AA;
 XX
 Alignment Scores:
 Pred. No.: 0 Length: 771
 Score: 4201.00 Matches: 771
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 85.80% Indels: 0
 DB: 23 Gaps: 0
 US-09-774-490-1 (1-2709) x ABG96413 (1-771)
 QY 200 ATGGGCTGGTAACTAGAGATTGCTGCTTTCTGGGGAGTATTACTTACAGCAGAGCA 259
 |||||||
 Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThraAlaArgAla 20
 QY AACTATCAGAAATGGAAGAACAAATGTGCCAAGGTGAAATTTATCTACAAAGAAATGTTG 319
 |||||||
 Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
 QY GAATCCAAATGTGATCATCTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCATACCTTC 379
 |||||||
 Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
 QY CTTTGGATGAGGAACGAGTAGGCTGTATCTTGGAGCAAGAGATCATATTTTCATTC 439
 |||||||
 Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisPheSerPhe 80
 QY GACCTGGTTAATCAAGGATTTTCAAAAGATTGTGTGCCAGCATCTTACACCAAGAGA 499
 |||||||
 Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
 QY GATGAATGCAAGTGGCTGGAAAGACATCTCTGAAAGATGTGTCTAATTTTCATCAAGTA 559
 |||||||
 Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
 QY CTTAAGGCATATATCAGACTCCTGTGACCTGTGGAAACGGGGCTTTTCATCCCAATT 619
 |||||||
 Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
 QY TGCACTACATGAAATTTGACATCATCTGAGACAAATATTTTAACTGGAGAACTCA 679
 |||||||
 Db 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
 QY CATTTTGAACGGCCGTGGGAAGAGTCCATATACCTTAAGCTGTCTGACAGCATCCCTT 739
 |||||||
 Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
 QY TTAATAGATGAGAAATATATCTCTGGAACCTGACCTGATTTTATGGCGGAGACTTCTCT 799
 |||||||
 Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
 QY ATCTTCGGAACCTCTGGCGCACCCACCAATCAGGACAGCAGCAGCATGATTCAGAGTGG 859
 |||||||
 Db 201 IlePheArgThrLeuGlyHisHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
 QY CTCAATGATCAAGTTCTATTAGTGGCCACCTCATCTCAGAGAGTGAACAATCTCTGAAGAT 919
 |||||||
 Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
 QY GACAAAGTATATCTTTCTCCGTAAGAAATCAATAGTAGAGACACTCTCTGGAAGAGCT 979
 |||||||
 Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
 QY ACTCACGCTAGATAGATGATATGCAAGATGACATTTTGAGGGCCACAGAGTCTGTGTG 1039
 |||||||
 Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
 QY AATAAATGCAACATTTCTCAAGCTGTCTGATTTGCTCAGTCCAGTCCAGGTCCTCAATGCG 1099
 |||||||
 Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
 QY ATTGACACTCATTTTGTGAACTGCGAGATGTATCTCTAATGACTTTTAAAGATCTCTAAA 1159
 |||||||
 Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
 QY AATCCAGTGTATATGAGTGTGTTTACGACTTCCAGTAACTATTTTCAAGGGATCAGCGGTG 1219
 |||||||
 Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
 QY TGTATGTATAGCATGATGTGTGAGAAAGGTGTCTCTCGTCCATATGCCACAGGAT 1279
 |||||||
 Db 341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
 QY GGACCCCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTTATCCACGCCAGGAACT 1339
 |||||||
 Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380

QY 1340 TGTCCAGCAAAACATTTGGTGGTTTGTACTCTACAAAGGACCTTCCCTGATGATCTTATA 1399
DB 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspValIle 400
QY 1400 ACCTTTGCAAGAGTCATCCAGCCATGATCAATCCAGTGTTCCTATGAACAATCGCCCA 1459
DB 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTATCAAAACGAGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1519
DB 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 CGAGAGATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1579
DB 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGAGAGACTTGGTATGATTTAGAGAGGTTTCTGCTGAAGAAATG 1639
DB 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuGluGluMet 480
QY 1640 ACAGTTTTCGGAACCGACTGCTATTTCCAGCAATGAGGCTTCCACTAAGCAGCAACA 1699
DB 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATTGTTCAACGCTGGGTTGCCAGCTCCCTTTACACGGGTGATATTTAC 1759
DB 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GCGAAGCGGTGCTGAGTGTGCTCGCCGAGACCCCTTACTGTGCTGGATGGTCT 1819
DB 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTCCCACTGCAAGAGACGCAAGCAGCAAGATATAGAAT 1879
DB 541 AlaCysSerArgTyrPheProThrAlaLysArgArgArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTGCTTACAGCTTACACCATGATATACCATGGCCACAGCCCT 1939
DB 561 GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisGlyHisSerPro 580
QY 1940 GAAGAGAGAAATCATCTATGTTAGAGATAGTAGACATTTTGGATGCGAGTCCGAG 1999
DB 581 GluGluArgIleIleTyrGlyValGluAsnSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGCCCTGCTATTTGGCAATTCAGAGCGCAATAGAGCGCAAAAGAGAG 2059
DB 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGTGTCTACAA 2119
DB 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGGATTCAGGCAATACCTCTGCCATGCGGTGGAACATGGTTCATACAACTCTT 2179
DB 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAGGTAAACCTGGAAGTCAATGACACAGAGCTTTGGAGAGACTTCTTCATTAAGAT 2239
DB 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
QY 2240 GATGATGAGATGGCTCTAAGACCAAGAAATGTCATAGCATGACACCTAGCCAGAG 2299
DB 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGGTACAGACTTCATGAGCTCATCAACCCCAATCTCAACACGATGGATGAG 2359
DB 701 ValTyrTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGACAGTTTGGAAAGGACCGCAACACGCTCGCAAGCCAGGACCATACC 2419
DB 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740

QY 2420 CCAGGGAACAGTAACAAATGGAAGCCTTCAAGAAATATAGAAAGTAGAAACAGGAGG 2479
DB 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCACGAATTTAGAGGGGACCCAGGAGTGTCT 2512
DB 761 ThrHisGluPheGluArgAlaProArgSerVal 771
RESULT 4
ID AAY21264 standard; Protein; 796 AA.
XX AAY21264;
DT 22-JUL-1999 (first entry)
XX Human semaphorin III wild type protein fragment 1.
DE Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
KW frameshift mutation; age-related disease; neurodegenerative disorder;
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
KW high mobility group protein-C; neuroendocrine specific protein A.
XX Homo sapiens.
OS WO9845322-A2.
PN 15-OCT-1998.
PD 02-APR-1998; 98WO-IB00705.
PF 10-APR-1997; 97US-0043163.
PR (UYUT-) RIJKSUNIV UTRECHT.
PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
PA (UYRO-) UNIV ROTTERDAM ERASMUS.
PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
XX WPI; 1998-609901/51.
DR N-PSDB; AAX75767.
XX Diagnosing disease by detecting frameshift mutations in RNA or
PT corresponding protein mutations - used to diagnose cancer and
PT neurological diseases, particularly Alzheimer's disease, and also
PT for treatment and prevention with specific ribozymes or wild-type
PT RNA
XX Disclosure; Figure 16; 258pp; English.
XX This invention describes a novel method for the diagnosis of a disease
CC caused by, or associated with, an RNA molecule that has a frameshift
CC mutation. The method is used to diagnose age-related diseases, especially
CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
CC and many others listed) or susceptibility to these disorders. The method
CC allows a definitive diagnosis of Alzheimer's disease in living patients,
CC at an early stage. It is based on the observation that disease may be
CC caused by mutations in RNA rather than DNA. The invention describes the
CC use of neuronal system RNA molecules, specifically proteins including
CC beta-amyloid precursor protein (beta-APP), the microtubule associated
CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group

CC protein-C (HMGP-C) and neuroendocrine specific protein A.

XX Sequence 796 AA;

Alignment Scores:

Pred. No.: 0 Length: 796
Score: 4194.00 Matches: 773
Percent Similarity: 97.48% Conservative: 0
Best Local Similarity: 97.48% Mismatches: 0
Query Match: 85.66% Indels: 20
DB: 19 Gaps: 1

US-09-774-490-1 (1-2709) x RAY21264 (1-796)

QY 194 TCGAGCTGGCTGGTAACTAGGATGCTGCTCTTTCTGGGAGTATTACTACAGCA 253
DB 4 CysSerMetGlyTTrpLeuThrArgileValCysLeuPheThrGlyValLeuLeuThra 23
QY 254 AGAGCAAACTATCAGAAATGGGAGAACAAATGTGCCAAGCTGMAATTTATCTACAAAGAA 313
DB 24 ArgAlaAsnTyrGlnAsnGlyLysAsnValProArgLeuLeuLeuSerTyrLysGlu 43
QY 314 ATGTTGGAAATCAACAATGTGATCACTTTCAATGCTTGGCCACAGCTCCAGTTATCAT 373
DB 44 MetLeuGluSerAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHis 63
QY 374 ACCTTCCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATATT 433
DB 64 ThrPheLeuLeuAspGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePhe 83
QY 434 TCATTGCACTGGTTAATATCAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACACC 493
DB 84 SerPheAspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThr 103
QY 494 AGAGAGATGAATCAAGTGGCTGGAAAGACATCTCTGAAGATGCTAAATTTTCATC 553
DB 104 ArgArgAspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPhelle 123
QY 554 AAGTACTTAAGGCATATAATCAGACTCACCTTGTACGCTGTGGAAACGGGGCTTTTCAT 613
DB 124 LysValLeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHis 143
QY 614 CCAATTCGACTACATTAATGGAATCATCATCTGAGGACATATTTTAAAGCTGGAG 673
DB 144 ProIleCysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGlu 163
QY 674 AACTCACATTTGAAACGGCGTGGGAGAGTCCATATGACCTTAAGCTCTGCACGCA 733
DB 164 AsnSerHisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAla 183
QY 734 TCCCTTTTAAATAGATGAGAAATTAATCTCTGGAACCTGCAGCTGATTTTATGGGCGAGAC 793
DB 184 SerLeuLeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAsp 203
QY 794 TTGCTATCTTCGGAACCTTTGGGACCAACCAATCAGGACAGACATGATTC 853
DB 204 PheAlaIlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSer 223
QY 854 AGGTGGCTCAATGATCCAAATTCATAGTCCCACTCATCTCAGAGAGTGCACATCT 913
DB 224 ArgTrpLeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnPro 243
QY 914 GAAGATGACAAAGATATATCTTTCTTCGTCGAAATGCAATAGATGGAGAACACTCTGGA 973
DB 244 GluAspAspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGly 263
QY 974 AAAGCTACTCGCTAGATAGGTGAGATATGCAAGATGATTTTGGAGGCCACAGAGT 1033
DB 264 LysAlaThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSer 283
QY 1034 CTGTTGAATAATGGACACATCTCTCAAGCTCGTCTGATTTGCTAGTCCAGGTCCA 1093
DB 284 LeuValAsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyPro 303

QY 1094 AATGGCATTGACACTCATTTTGTATGAAGTGCAGGATGTATTCTTAATGAACCTTTAAAGAT 1153
DB AsnGlyIleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAsp 323
QY 1154 CCTAAATCCAGTTGTATATATGAGTGTATACGATCTCCAGTAACATTTTCAAGGATCA 1213
DB ProLysAsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySer 343
QY 1214 GCCGTGTGTATAGCATGATGATGTGAGAAGGGTGTCTTGGTCCATATGCCCCAC 1273
DB AlaValCysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHis 363
QY 1274 AGGATGGACCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCACGGCCA 1333
DB ArgAspGlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgPro 383
QY 1334 GGAATCTGTCCACACAAACATTTGGTGGTGTGACTCTACAAAGGACCTTCTGTGATCAT 1393
DB GlyThrCysProSerLysThrPheGlyLysAspSerThrLysAspLeuProAspAsp 403
QY 1394 GTTATAACCTTTGCAAGAAGTCAATCCAGCCATGACATCCAGTGTCTTCTCT - 1444
DB ValIleThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsn 423
QY 1445 -----ATGAACAAT 1453
DB ArgProIleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValMetAsnAsn 443
QY 1454 CGCCCAATAGTGATCAAAACGGATGTAATATCAATTTACACAAATGTCTAGACCGA 1513
DB ArgProIleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArg 463
QY 1514 GTGATGTCAGAAAGATGACAGTATGATGTTATCGGAACAGATGTTGGGACCGTT 1573
DB ValAspAlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrVal 483
QY 1574 CTTAAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAGAGGTTTCTGCTGAA 1633
DB LeuLysValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGlu 503
QY 1634 GAAATGACAGTTTTCGGGAACCGACTGCTATTTTACAGCAATGGAGCTTTCACATAGCAG 1693
DB GluMetThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGln 523
QY 1694 CAACAACTATATATTGTTCAACGGCTGGGCTGCCAGCTCCCTTTACACCGGTGTGAT 1753
DB GlnGlnLeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAsp 543
QY 1754 ATTTACGGGAAGCGTGTGCTGAGTGTGCTCCTCCCGAGACCCCTTACTGTGCTTGGAT 1813
DB IleTyrGlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAsp 563
QY 1814 GGTTCGCTGCTTCTCCTATTTTCCACTGCAAGAGACGCAACAGACGACAGATATA 1873
DB GlySerAlaCysSerArgTyrPheProThrAlaLysArgArgThrArgGlnAspIle 583
QY 1874 AGAATGGAGACCCACTGACTCTGACTGTTACACATGATTAATCACATGGCCAC 1933
DB ArgAsnGlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisIleGlyHis 603
QY 1934 AGCCTCGAAGAGAAATCATCTATGGTGTAGAGAAATAGTAGCACATTTTTCGAATGCACT 1993
DB SerProGluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSer 623
QY 1994 CCGAAGTCGAGAGAGCGCTGCTTATTTGGCAATTCAGAGCGCAATTAAGAGCGCAAAA 2053
DB ProLysSerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLys 643
QY 2054 GAAGAGATCAGCTGGATGATCATATCATCAGGACAGATCAAGCCCTTCTGCTAGGTAGT 2113
DB GluGluIleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSer 663

QY 2114 CTACACAGAGGATTTCAGGCAATACCTCTCCATGGGTGGAACATGGGTTCATACAA 2173
 DB 664 LeuGlnGlnLysAspSerGlyAsnTrpLeuGlnHisAlaValGluHisGlyPheIleGln 683
 QY 2174 ACTCTCTTAAGGTAACCTCGAAGTCATTGTGACAGAGCAATTTGGAAGAATCTTCTCAT 2233
 DB 684 ThrLeuLeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHis 703
 QY 2234 AAGATGATGATGAGATGGCTCTAAGACCAAGAAATGTCCAATAGATGACACCTAGC 2293
 DB 704 LysAspAspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSer 723
 QY 2294 CAGAAGGTCCTGTACAGAGACTTCATGAGCTCATCAACCCCAATCTCAACAGCATG 2353
 DB 724 GlnLysValTrpTrpArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMet 743
 QY 2354 GATGAGTCTGTGAACAAAGTTTGGAAAAGGGACCGAAAACAAACGTCGGCAAGGCCAGGA 2413
 DB 744 AspGluPheCysGluGlnValTrpLysArgAspArgLysGlnArgProGly 763
 QY 2414 CATACCCAGGGAACAGTACAAATGGAAGCACTTACAAGAAATAAGAAAGGTAGAAC 2473
 DB 764 HisThrProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsn 783
 QY 2474 AGGAGGACCCAGCAATTTGAGAGGGCCACCCAGGAGTGTC 2512
 DB 784 ArgArgThrHisGluPheGluArgAlaProArgSerVal 796

RESULT 5

AAR74175
 ID AAR74175 standard; Protein; 477 AA.

XX AC AAR74175;

XX DT 01-NOV-1995 (first entry)

XX DE Human collapsin.

XX KW Collapsin; antibody; therapy.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Binding-site 9..19 /note= "antibody binding site"

FT Binding-site 51..65 /note= "antibody binding site"

FT US5416197-A.

XX PD 16-MAY-1995.

XX PF 15-OCT-1993; 93US-0136922.

XX PR 15-OCT-1993; 93US-0136922.

XX PA (UYPE-) UNIV PENNSYLVANIA.

XX PI Luo Y, Raper JA;

XX DR WPI; 1995-193478/25.

XX DR N-PSDB; AAQ92331.

XX PT New antibody to human collapsin - used to inhibit the activity of
 PT collapsin, to induce neurite out-growth and to treat individuals with
 PT nerve damage.

XX PS Claim 2; Columns 15-18; 11pp; English.

XX CC Aff antibody capable of specifically binding at least a portion of
 CC the collapsin protein can be used to purify human collapsin and
 CC to inhibit the activity of the protein. It can be used to induce
 CC neurite outgrowth by neuronal cells and to treat individuals

CC suffering from nerve damage.

XX SQ Sequence 477 AA;

Alignment Scores:

Pred. No.: 1.18e-227 Length: 477
 Score: 2562.00 Matches: 472
 Percent Similarity: 99.37% Conservative: 0
 Best Local Similarity: 99.37% Mismatches: 3
 Query Match: 52.33% Indels: 0
 DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x AAR74175 (1-477)

QY 899 GAGAGTCACATCTCTGAGATGACAAAGTATACCTTTCTCCGTGAAAATGCAATAGAT 958
 DB 2 GluHisAspAsnProGluAspAspLysValTrpPhePheArgGluAsnAlaIleAsp 21
 QY 959 GGAGAACACTCTGGAAAAGCTACTCAGCTAGATAAGTACAGATATGCAAGAATGACTTT 1018
 DB 22 GlyGluHisSerGlyLysAlaThrHisAlaArgIleGlyGlnIleCysLysAsnAspPhe 41
 QY 1019 GGAGGGCACAGAACTCGTGAATAAATGGACAAACATTCCTCAAGCTCGTCTGATTTGC 1078
 DB 42 GlyGlyHisArgSerLeuValAsnLysTrpThrThrPheLeuLysAlaArgLeuIleCys 61
 QY 1079 TCAGTGGCAGGTCCAAATGGCATTGACACTCATTTTGTATGTAAGTCAAGTCAAGGATGATTCCTA 1138
 DB 62 SerValProGlyProAsnGlyIleAspThrHisPheAspGluLeuGlnAspValPheLeu 81
 QY 1139 ATGAACCTTTAAAGATCCTAAAAATCCAGTTGTATATGAGTGTATTTACGACTTCCAGTAAC 1198
 DB 82 MetAsnPhelysAspProLysAsnProValValTrpGlyValPheThrThrSerSerAsn 101
 QY 1199 ATTTTCAAGGATCAGCGTGTGTATGTATAGCATGATGTAGTGTAGAGAGGTGTCTT 1258
 DB 102 IlePheLysGlySerAlaValCysMetTrpSerMetSerAspValArgValPheLeu 121
 QY 1259 GGTCCATATGCCACAGGATGGACCCCACTATCAATGGGTGCCTTATCAAGGAGAGATC 1318
 DB 122 GlyProTrpAlaHisArgAspGlyProAsnTrpGlnTrpValProTrpGlnGlyArgVal 141
 QY 1319 CCTATCCAGCCGAGCAACTGTCCAGCAAAAACATTTGGTGGTGTGTGTACTTACAAAG 1378
 DB 142 ProTrpProArgProGlyThrCysProSerLysThrPheGlyGlyPheAspSerThrLys 161
 QY 1379 GACTTCTGTATGATTTATTAACCTTTGCAAGAAGTCAATCCAGCATGTACATCCAGTG 1438
 DB 162 AspLeuProAspAspValIleThrPheAlaArgSerHisProAlaMetTrpAsnProVal 181
 QY 1439 TTTCCTATGACAAATCGCCCAATAGTATCAAAACGGATGTAATTTATCAATTTACACAA 1498
 DB 182 PheProMetAsnAsnArgProIleValIleLysThrAspValAsnTrpGlnPheThrGln 201
 QY 1499 ATTGTCGTAGACCGAGTGGATGCAAGAAGTATGATGTATGTATGTTTATCGGAACA 1558
 DB 202 IleValValAspArgValAspAlaGluAspGlyGlnTrpAspValMetPheIleGlyThr 221
 QY 1559 GATGTGGACCGTCTTTAAAGTAGTTTCAATTCCTAAGGAGACTGTGTATGATTTAGAA 1618
 DB 222 AspValGlyThrValLeuLysValValSerIleProLysGluThrTrpTrpAspLeuGlu 241
 QY 1619 GAGTTCCTGCTGGAGAAATGACAGTTTTCGGGAACCGACTGTATTTACAGCAATGAG 1678
 DB 242 GluValLeuLeuGluGluMetThrValPheArgGluProThrAlaIleSerAlaMetGlu 261
 QY 1679 CTTTCCACTAAGCAGCAACACTATATATTGTTTCAACGGTGGGGTGTCCAGCTCCCT 1738
 DB 262 LeuSerThrLysGlnGlnLeuTrpIleGlySerThrAlaGlyValAlaGlnLeuPro 281
 QY 1739 TTACACCGGTGTGATATTTACGGGAACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1798
 DB 282 LeuHisArgCysAspIleTrpGlyLysAlaCysAlaGluCysCysLeuAlaArgAspPro 301

QY 1799 TACTGTGCTGGATGGTTCTGCATGTTCTCGCTATTTTCCACTGCAAGAGAGCGACA 1858
 Db |||||
 QY 302 TyrCysAlaTrpAspGlySerAlaCysSerArgTyrPheProThrAlaLysArgArgThr 321
 Db |||||
 QY 1859 AGACACCAAGATATAAGAAATGGACACCCACTGACTCTCTGCTTACAGCTTACACCATGAT 1918
 Db |||||
 QY 322 ArgArgGlnAspIleArgAsnGlyAspProLeuThrHisCysSerAspLeuHisHisAsp 341
 Db |||||
 QY 1919 AATCACCATGGCCACAGCCCTGAAGAGAGATCATCTATGGTGTAGAGAAATAGTAGACA 1978
 Db |||||
 QY 342 AsnHisHisGlyHisSerProGluGluArgIleIleTyrGlyValGluAsnSerSerThr 361
 Db |||||
 QY 1979 TTTTGGAAATGACGATCCGAGTCGAGAGAGCGCTGTCTATTGGCAATTCAGAGGCGA 2038
 Db |||||
 QY 362 PheLeuGluCysSerProLysSerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArg 381
 Db |||||
 QY 2039 AATGAAGAGCGAAAGAGAGATCAGAGTGCAGTGCATCATCATCAGCAGACATCAAGC 2098
 Db |||||
 QY 382 AsnGluGluArgLysGluGluIleArgValAspHisIleIleArgThrAspGlnGly 401
 Db |||||
 QY 2099 CTTCTGCTACGTAGTCTCAACAGAGAGATTCAGGCAATTCAGGCTGCGGTGGA 2158
 Db |||||
 QY 402 LeuLeuLeuArgSerLeuGlnGlnLysAspSerGlyAsnTyrLeuCysHisAlaValGlu 421
 Db |||||
 QY 2159 CATGGTTTCATACAACTCTTCTTAAGGTAAACCTGGAAGTCAATGCACAGAGCATTTG 2218
 Db |||||
 QY 422 HisGlyPheIleGlnThrLeuLeuLysValThrLeuGluValIleAspAsnGluHisLeu 441
 Db |||||
 QY 2219 GAAGAACTCTTCATTAAGATGATGATGAGATGCTCTTAAGCAAAAGAAATGTCGAT 2278
 Db |||||
 QY 442 GluGluLeuLeuHisLysAspAspGlyHisGlySerLysThrLysGluMetSerAsn 461
 Db |||||
 QY 2279 AGCATGACACCTAGCCAGAAGTCTGTGACAGAGACTTTCATGCGAG 2323
 Db |||||
 QY 462 SerMetThrProSerGlnLysValTrpTyrArgAspPheMetGln 476
 Db |||||

RESULT 6

AAV27127
 ID AAV27127 standard; Protein; 777 AA.
 XX
 AC AAV27127;

DT 14-SEP-1999 (first entry)
 XX
 DE Human brain tissue-derived polypeptide (clone OM007).

XX Brain tissue; human; bone marrow; umbilical cord venous endothelial cell;
 KW recombinant; diagnosis; treatment.
 XX

OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..36

FT Protein /note= "signal peptide"

FT Protein 37..777

FT Protein /note= "mature protein"

XX WO9933873-A1.

XX 08-JUL-1999.

XX 25-DEC-1998; 98WO-JP05952.

XX 26-DEC-1997; 97JP-0358811.

XX (ONÖY) ONO PHARM CO LTD.

XX Fukushima D, Shibayama S, Tada H;

XX WPI; 1999-419088/35.

XX N-PSDB; AAX89112, AAX89113.

XX

PT New adult human brain tissue-produced polypeptides useful for
 diagnosis and treatment

PS Claim 1; Page 36-39; 86pp; Japanese.

CC The invention provides polypeptides (AAV27127-Y27133) produced by human
 adult brain tissue, human bone marrow or a human umbilical cord venous
 endothelial cell. Host cells transformed with vectors comprising the
 nucleic acids encoding the polypeptides are used for the recombinant
 expression of the polypeptides. The polypeptides can be used in
 diagnosis, treatment and basic studies, with wide applications in
 treatment depending on the activity to be aimed at. Sequences
 CC AAX89112-125 represent nucleic acids encoding the polypeptides.

XX SQ Sequence 777 AA;

Alignment Scores:

Pred. No.: 2,68e-198 Length: 777
 Score: 2245.50 Matches: 413
 Percent Similarity: 73.67% Conservative: 141
 Best Local Similarity: 54.92% Mismatches: 177
 Query Match: 145.86% Indels: 21
 DB: 20 Gaps: 9

US-09-774-490-1 (1-2709) x AAV27127 (1-777)

QY 275 AAGAACAAATGTCACCAAGGCTGAAATATCTCTACAAAGAAATGTTGGAATCCAAACAATGTG 334
 Db |||||
 QY 39 LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuSerAsnSerCys 58
 Db |||||
 QY 335 ATCACTTTCATGCTTGGCCACAGCTCCAGTTATCATATCATCTCTCTCTTGGATGAGGAA 394
 Db |||||
 QY 59 IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuLeuAspGluGlu 78
 Db |||||
 QY 395 CGGAGTAGGCTGTATGTTGGAGCAAGATCACATATTTTCATCGACCTGGTAAATATC 454
 Db |||||
 QY 79 ArgGlyArgLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
 Db |||||
 QY 455 ---AAGGATTTTCAAAAGATTTGTGCGCATCTTACACCAAGAGAGATGAAATGCAAG 511
 Db |||||
 QY 99 AsnLysAsnPheLysLysIleTyrTrpProAlaLysGluArgValGluLeuCysLys 118
 Db |||||
 QY 512 TGGCTCGAAAGACATCTCTGAAAGATGTCTAAATTTTCATCAAGGTACTTAAAGCATAT 571
 Db |||||
 QY 119 LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr 138
 Db |||||
 QY 572 AATCAGACTCACTGTAGCTGTGGAAACGGGGCTTTTCATCCAAATTTGACCTACATTT 631
 Db |||||
 QY 139 AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle 158
 Db |||||
 QY 632 GAAATTCGACATCATCTGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTTGAAGAC 691
 Db |||||
 QY 159 AspLeuGlyValTyrLysGluAspIleIlePheLysLeuAspThrArgAsnLeuGluSer 178
 Db |||||
 QY 692 GGCCGTGGAGAGATCCATATGACCCCTAAGCTGTGACAGCATCCCTTTTAAATAGATGGA 751
 Db |||||
 QY 179 GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu 198
 Db |||||
 QY 752 GAATATACCTGGAAGTCTGAGCTGATTTTATGGGCGAGACTTTGCTATCTTCGGAAC 811
 Db |||||
 QY 199 TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer 218
 Db |||||
 QY 812 CTTGGG-----CACCACCCCAATCAGGACAGCAGCATGATTCAGGTGGCTC 862
 Db |||||
 QY 219 LeuGlyProThrHisAspHisHisTyrIleArgThrAspIleSerGluHisTyrTrpLeu 238
 Db |||||
 QY 863 AATGATCCAAAGTTCATTAGTGGCCACCTCATCTCAGAGAGTGAACATCTCTGAAGATGAC 922
 Db |||||
 QY 239 AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrAsnProAspAsp 258
 Db |||||
 QY 923 AAGTATACCTTTTCTTCCTGAGAAATGCAATAGATGAGAACACTCTCTGGAAGAGTACT 982
 Db |||||
 QY 259 LysIleTyrPhePheArgGluSerSerGlnGluGlySerThrSerAspLysThrIle 278
 Db |||||

QY 983 CAGCTAGAAAGTGCAGATATGCAAGAATGACTTTGGAGGACACAGAGTCTGGTGAAT 1042
Db
QY 279 LeuSerArgValGlyArgValCysLysAsnAspValGlyGlnArgSerLeuLeuIleAsn 298
Db
QY 1043 AAATGACAACTTCCTCAAGCTCGCTGATTGCTCAGTGCAGGTCCTCAAAATGGCAATT 1102
Db
QY 299 LysTrpThrPheLeuLysAlaArgLeuIleCysSerIleProGlySerAspGlyAla 318
Db
QY 1103 GACACTCATTTGATGAAGTGCAGGATGATATTCCTTAATCACTTTAAAGATCCTAAAT 1162
Db
QY 319 AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArgAsn 338
Db
QY 1163 CAGTTGTATATGAGTGTTCAGCTTCAGTACATTTCAAGGATCCAGCTGTGT 1222
Db
QY 339 ProValValTyrGlyValPheThrThrThrSerSerIlePheLysGlySerAlaValCys 358
Db
QY 1223 ATGTATAGCATGATGTGAGAGGTTCTTCCTTGGTCCATATGCCACAGGATGGA 1282
Db
QY 359 ValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysGluSer 378
Db
QY 1283 CCAACTATCAATGGTGCTTATCAAGAGAGTCCCTTATCCAGCGCCAGAACTGT 1342
Db
QY 379 AlaAspHisArgTrpValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys 398
Db
QY 1343 CCAGCAAAACATTTGGTGT---TTTGACTCTCAAAAGGACTTCCTGATGATGTATA 1399
Db
QY 399 ProSerLysThrTyrAspProLeuLysSerThrThrArgAspPheProAspValIle 418
Db
QY 1400 ACCTTTGCAAGAAGTCAATCCAGCTGTATCAATCCAGTGTTCCTATGAACATCCCA 1459
Db
QY 419 SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro 438
Db
QY 1460 ATAGTATCAAAACGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1519
Db
QY 439 ThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIle 458
Db
QY 1520 CGAGAGATGGACATGATGTTATGTTTATCGGAACAGATGTTGGGACCGTCTTAA 1579
Db
QY 459 AlaGluAspGlyGlnThrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
Db
QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGTGTATGATTTAGAGAGGTTCTGCTGGAAGAAATG 1639
Db
QY 479 ValValSerIleSerLysGluLysTrp---AsnMetGluGluValValLeuGluGluLeu 497
Db
QY 1640 ACAGTTTTCGGGAACCGACTGCTATTCAGCAATGGAGCTTTCACATGAAGCAGCAACA 1699
Db
QY 498 GlnIlePheLysHisSerSerIleLeuAsnMetGluLeuSerLeuLysGlnGlnGln 517
Db
QY 1700 CTATATATGTTTCAACGGCTGGGTGCTCCAGCTCCCTTTACACGGTGTATATTAC 1759
Db
QY 518 LeuTyrIleLysSerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
Db
QY 1760 GGGAAAGCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTTGGATGTTCT 1819
Db
QY 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAsn 557
Db
QY 1820 GCATGTTCTCGCTATTTTCCCTGCAAGAGAGCAGCAGCAGCAGATATAGAAT 1879
Db
QY 558 AlaCysSerArgTyrAlaProThrSerLysArgAlaArgGlnAspValLysTyr 577
Db
QY 1880 GGAGACCCACTCACTCACTGTTCCAGATTCACCATGATATACCATGATCAATCAGCCGACCCCT 1939
Db
QY 578 GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GlnThrAla 596
Db
QY 1940 GAAGAGAGATCATCTATGCTGTAGAGATAGTACACATTTTGAATGCAAGTCCGAG 1999
Db
QY 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616
Db
QY 2000 TGCAGAGAGCGCTGTCTATTGGCAATTCAGAGCGGCAATGAAGAGCGCAAAAGAG 2059
Db
QY 617 SerGlnGlnAlaThrIleLysTrpTyrIleGlnArgSerGlyAspGluHisArgGluGlu 636
Db

QY 2060 ATCAGATGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGTAGTACAA 2119
Db
QY 637 LeuLysProAspGluArgIleIleLysThrGlnTyrGlyLeuLeuIleArgSerLeuGln 656
Db
QY 2120 CAGAAGGATTCAGCAATTAACCTCTGCGCATCGGTGGAAACATGGGTTTCATCAAACTCTT 2179
Db
QY 657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGlnHisThrPheIleHisThrIle 676
Db
QY 2180 CTTAAGGTAAACCTCGAAGTCAATTCACACAGAGATTTGGGAAGAACTTCTTCATAAAGAT 2239
Db
QY 677 ValLysLeuThrLeuAsnValIleGluAsnGlnGlnMetGluAsnThrGlnArgAlaGlu 696
Db
QY 2240 GATGATGAGATGCTCTTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAGAG 2299
Db
QY 697 HisGlnGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
Db
QY 2300 GTCTGTACAGACTTTCATCAGCTCATCAACACCCCAATCTCAACACAGATGATGAG 2359
Db
QY 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln 729
Db
QY 2360 TTCTGTGACAACTTTGGAAAGGACCGCAAAACACGTCGCAAGGCCAGGACATACC 2419
Db
QY 730 TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly----- 747
Db
QY 2420 CCAGGAAACAGTAAACAAATGGAAGCACTTCAAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db
QY 748 -----ProLysTrpLysHisMetGlnGluMetLysLysLysArgAsnArgArg 763
Db
QY 2480 ACCCAC---GAATTTGAGAGGCGCCAGGAGTGTCT 2512
Db
QY 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775
Db
RESULT 7
AAI99427
ID AAY99427 standard; Protein; 777 AA.
XX AC AAY99427;
XX DT 08-AUG-2000 (first entry)
XX DE Human PRO1491 (UNQ760) amino acid sequence SEQ ID NO:310.
XX KW Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;
XX KW transmembrane; secretion; immunoadhesion; pharmaceutical; screening.
XX OS Homo sapiens.
XX PN WO200012708-A2.
XX PD 09-MAR-2000.
XX PF 01-SEP-1999; 99WO-US20111.
XX PR 01-SEP-1998; 98US-0098716.
XX PR 01-SEP-1998; 98US-0098749.
XX PR 01-SEP-1998; 98US-0098750.
XX PR 02-SEP-1998; 98US-0098803.
XX PR 02-SEP-1998; 98US-0098821.
XX PR 02-SEP-1998; 98US-0098843.
XX PR 09-SEP-1998; 98US-0099536.
XX PR 09-SEP-1998; 98US-0099596.
XX PR 09-SEP-1998; 98US-0099598.
XX PR 09-SEP-1998; 98US-0099602.
XX PR 09-SEP-1998; 98US-0099642.
XX PR 10-SEP-1998; 98US-0099741.
XX PR 10-SEP-1998; 98US-0099754.
XX PR 10-SEP-1998; 98US-0099763.
XX PR 10-SEP-1998; 98US-0099792.
XX PR 10-SEP-1998; 98US-0099808.
XX PR 10-SEP-1998; 98US-0099812.
XX PR 10-SEP-1998; 98US-0099815.
XX PR 10-SEP-1998; 98US-0099816.
XX PR 15-SEP-1998; 98US-0100385.


```
PR 15-SEP-1998; 98US-0100388.
PR 15-SEP-1998; 98US-0100390.
PR 16-SEP-1998; 98US-0100584.
PR 16-SEP-1998; 98US-0100627.
PR 16-SEP-1998; 98US-0100661.
PR 16-SEP-1998; 98US-0100662.
PR 16-SEP-1998; 98US-0100664.
PR 17-SEP-1998; 98US-0100683.
PR 17-SEP-1998; 98US-0100684.
PR 17-SEP-1998; 98US-0100710.
PR 17-SEP-1998; 98US-0100711.
PR 17-SEP-1998; 98US-0100919.
PR 18-SEP-1998; 98US-0100930.
PR 18-SEP-1998; 98US-0100848.
PR 18-SEP-1998; 98US-0100849.
PR 18-SEP-1998; 98US-0101014.
PR 18-SEP-1998; 98US-0101068.
PR 18-SEP-1998; 98US-0101071.
PR 22-SEP-1998; 98US-0101279.
PR 23-SEP-1998; 98US-0101471.
PR 23-SEP-1998; 98US-0101472.
PR 23-SEP-1998; 98US-0101474.
PR 23-SEP-1998; 98US-0101475.
PR 23-SEP-1998; 98US-0101476.
PR 23-SEP-1998; 98US-0101477.
PR 24-SEP-1998; 98US-0101479.
PR 24-SEP-1998; 98US-0101738.
PR 24-SEP-1998; 98US-0101741.
PR 24-SEP-1998; 98US-0101743.
PR 24-SEP-1998; 98US-0101915.
PR 24-SEP-1998; 98US-0101916.
PR 29-SEP-1998; 98US-0102207.
PR 29-SEP-1998; 98US-0102240.
PR 29-SEP-1998; 98US-0102307.
PR 29-SEP-1998; 98US-0102330.
PR 29-SEP-1998; 98US-0102331.
PR 30-SEP-1998; 98US-0102484.
PR 30-SEP-1998; 98US-0102487.
PR 30-SEP-1998; 98US-0102570.
PR 30-SEP-1998; 98US-0102571.
PR 01-OCT-1998; 98US-0102684.
PR 01-OCT-1998; 98US-0102687.
PR 02-OCT-1998; 98US-0102965.
PR 06-OCT-1998; 98US-0103258.
PR 06-OCT-1998; 98US-0103449.
PR 07-OCT-1998; 98US-0103344.
PR 07-OCT-1998; 98US-0103315.
PR 07-OCT-1998; 98US-0103328.
PR 07-OCT-1998; 98US-0103328.
PR 07-OCT-1998; 98US-0103396.
PR 07-OCT-1998; 98US-0103401.
PR 08-OCT-1998; 98US-0103633.
PR 08-OCT-1998; 98US-0103678.
PR 08-OCT-1998; 98US-0103679.
PR 08-OCT-1998; 98US-0103711.
PR 14-OCT-1998; 98US-0104257.
PR 20-OCT-1998; 98US-0104987.
PR 20-OCT-1998; 98US-0105000.
PR 20-OCT-1998; 98US-0105002.
PR 21-OCT-1998; 98US-0105104.
PR 22-OCT-1998; 98US-0105169.
PR 22-OCT-1998; 98US-0105266.
PR 26-OCT-1998; 98US-0105693.
PR 26-OCT-1998; 98US-0105694.
PR 27-OCT-1998; 98US-0105807.
PR 27-OCT-1998; 98US-0105881.
PR 27-OCT-1998; 98US-0105882.
PR 27-OCT-1998; 98US-0106062.
PR 28-OCT-1998; 98US-0106023.
PR 28-OCT-1998; 98US-0106029.
PR 28-OCT-1998; 98US-0106030.
PR 28-OCT-1998; 98US-0106032.
PR 28-OCT-1998; 98US-0106033.

PR 28-OCT-1998; 98US-0106178.
PR 29-OCT-1998; 98US-0106248.
PR 29-OCT-1998; 98US-0106384.
PR 29-OCT-1998; 98US-0106384.
PR 30-OCT-1998; 98US-0106500.
PR 30-OCT-1998; 98US-0106464.
PR 03-NOV-1998; 98US-0106856.
PR 03-NOV-1998; 98US-0106856.
PR 03-NOV-1998; 98US-0106902.
PR 03-NOV-1998; 98US-0106905.
PR 03-NOV-1998; 98US-0106919.
PR 03-NOV-1998; 98US-0106932.
PR 03-NOV-1998; 98US-0106934.
PR 10-NOV-1998; 98US-0107783.
PR 17-NOV-1998; 98US-0108775.
PR 17-NOV-1998; 98US-0108779.
PR 17-NOV-1998; 98US-0108787.
PR 17-NOV-1998; 98US-0108788.
PR 17-NOV-1998; 98US-0108801.
PR 17-NOV-1998; 98US-0108802.
PR 17-NOV-1998; 98US-0108806.
PR 17-NOV-1998; 98US-0108807.
PR 17-NOV-1998; 98US-0108867.
PR 17-NOV-1998; 98US-0108925.
PR 18-NOV-1998; 98US-0108848.
PR 18-NOV-1998; 98US-0108849.
PR 18-NOV-1998; 98US-0108850.
PR 18-NOV-1998; 98US-0108851.
PR 18-NOV-1998; 98US-0108852.
PR 18-NOV-1998; 98US-0108858.
PR 18-NOV-1998; 98US-0108904.
XX
XX (GETH ) GENENTECH INC.
XX
XX Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;
XX WPI: 2000-237871/20.
XX N-PSDB; AAA37109.
XX
XX New mammalian DNA sequences encoding transmembrane, receptor or
XX secreted PRO polypeptides, useful for screening of potential peptide or
XX small molecule inhibitors of the relevant receptor/ligand interactions
XX
XX Claim 12; Fig 176; 773pp; English.
XX
XX AAA37022 to AAA37144 encode the new isolated human transmembrane,
XX receptor or secreted PRO polypeptides given in AAY99340 to AAY99462. The
XX transmembrane and receptor PRO proteins can be used for screening of
XX potential peptide or small molecule inhibitors of the relevant
XX receptor/ligand interactions. The polypeptides and nucleotide sequences
XX encoding then have various industrial applications, including uses as
XX pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent
XX PCR primers and hybridisation probes used in the isolation of the PRO
XX polypeptides from the present invention.
XX
XX SQ Sequence 777 AA;

Alignment Scores:
Pred. No.: 2,68e-198 Length: 777
Score: 2245.50 Matches: 413
Percent Similarity: 73.67% Conservative: 141
Best Local Similarity: 54.92% Mismatches: 177
Query Match: 45.86% Indels: 21
DB: 21 Gaps: 9

US-09-774-490-1 (1-2709) x AAY99427 (1-777)

Qy 275 AAGAACAAATGTGCAAGCTGAATATCTACAAAGAAATGTTGGATCAACAATGTG 334
Db 39 LysGlnAsnIleProArgLeuLysLeuThrTyrIysAspLeuLeuLeuSerAsnSerCys 58
Qy 335 ATCATTTCATGGCTTGGCCACACAGCTCCAGTTATCATCTCTCTTTGGATGAGAA 394
Db 59 IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuLeuAspGluGlu 78
```

Qy	395	CGGAGTAGCGTGATGTTGGAGCAAGGATCA	CACATATTTTCATCTCGACCTGGTTATATC	454
Db	79	ArgGlyArgLeuLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu	98	
Qy	455	---AAGATTTTCAAAAGATTTGTGTGCCAGTATCTTACACCCAGNAGAGATGAATCGAG	511	
Db	99	AsnLysAsnPheIlystIleTyrTrpProAlaAlaLysGluArgValGluLeuCysLys	118	
Qy	512	TGGCTCGAAAAGACATCCTGMAAGATGTGCTAAATTTTCATCAAGGTACTTAAAGGCAT	571	
Db	119	LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr	138	
Qy	572	AATCAGACTCACTTGATAGCCTGTGGAAACGGGGCTTTTCATCAATTTGCACCTACATT	631	
Db	139	AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle	158	
Qy	632	GAATTTGGACATCATCTCTGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTTGA	691	
Db	159	AspLeuGlyValTyrLysGluAspIleIlePheLysLeuAspThrHisAsnLeuGluSer	178	
Qy	692	GGCCGTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAAATAGATGGA	751	
Db	179	GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu	198	
Qy	752	GAATTATATCTCGAACTGCAGCTGATTTTATGGCGCAGACTTTGCTATCTTCCGA	811	
Db	199	TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer	218	
Qy	812	CTTGGG-----CACCAACCACCAATCAGNACAGACAGCATGATTTCCAGTGGCTC	862	
Db	219	LeuGlyProThrHisAspHisHisTyrIleArgThrAspIleSerGluHisTyrTrpLeu	238	
Qy	863	AATGATCCAAAGTTCAATTAGTGCCCACTCATCTCAGAGAGTGACATCTCGAAGATGAC	922	
Db	239	AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrAsnProAspAspAsp	258	
Qy	923	AAAGTATATCTTTTCTCCGTGAAAATGCAATAGATGGAGAACA	982	
Db	259	LysIleTyrPhePhePheArgGluSerSerGlnGluGlySerThrSerAspLysThrIle	278	
Qy	983	CACGCTAGAATAGTCAAGATATGCAAGAATGACCTTTGGAGGCGACAGAACTCTGGTGA	1042	
Db	279	LeuSerArgValGlyArgValCysLysAsnAspValGlyGlyGlnArgSerLeuIleAsn	298	
Qy	1043	AAATGGACAATCTCTCAAGCTCGNCTGATTTGCTCAGTGCAGCTGCCAATGGCATT	1102	
Db	299	LysTrpThrThrPheLeuLysAlaArgLeuIleCysSerIleProGlySerAspGlyAla	318	
Qy	1103	GACACTATTTTGATGAACCTGCAGATGTATTCCTAATGAACCTTTAAGATCCCTAAA	1162	
Db	319	AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArgAsn	338	
Qy	1163	CCAGTTGTATAGGAGTGTTTACGACTTCGATGAACATTTTCAAGGATCAGCGGTGTGT	1222	
Db	339	ProValValTyrGlyValPheThrThrSerSerIlePheLysGlySerAlaValCys	358	
Qy	1223	ATGTATAGCATGACTGTGAGAGGGTGTTCCTTCCTCCATATGCCACAGGATGGA	1282	
Db	359	ValTyrSerMetAlaAspIleIleArgAlaValPheAsnGlyProTyrAlaHisLysGluSer	378	
Qy	1283	CCCAACTATCAATGGGTGCCTTATCAAGGAAGAGTGCCCTTATCCAGCGGCAGAACTGT	1342	
Db	379	AlaAspHisArgTrpValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys	398	
Qy	1343	CCGAGCAAAACATTTGGTGGT---TTTGACTCTCAAAAGGACCTTCTGTATGATGTATA	1399	
Db	399	ProSerLysThrTyrAspProLeuIleLysSerThrArgAspPheProAspAspValIle	418	
Qy	1400	ACCTTTGCAGAAGTCAATCCAGCCATGTACAAATCCAGTGTTCCTATGAACAATCCGCCA	1459	
Db	419	SerPheLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro	438	
Qy	1469	ATAGTGATCAAAACCGGATGTAAATATATCAATTTACAGAAATTTGCTGAGACCGAGTGGAT	1519	

Db	439	ThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIle	458
Qy	1520	GCAGACATGCACAGATGATGATGTTATGTCGGAACAGATGTCGGGACCGTCTTTAAA	1579
Db	459	AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys	478
Qy	1580	GTAGTTTCAATTCTTAAGGAGACTTGGTATGATTTAGAACAGGTTCTGCTGGGAAGAAATG	1639
Db	479	ValValSerIleSerLysGluLysTrp--AsnMetGluGluValValLeuGluGluLeu	497
Qy	1640	ACAGTTTTTCGGAAACCGACTGCTATTTCAGCAATCGAGCTTCCACTAAGCAGCAACAA	1699
Db	498	GlnIlePheLysHisSerIleIleLeuAsnMetGluLeuSerLeuLysGlnGlnGln	517
Qy	1700	CTATATATGGTTTCAACGGCTGGGGTTCGCCAGCTCCCTTTACACCGGTGTGATATTTAC	1759
Db	518	LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr	537
Qy	1760	GGGAAACGCTGCTCGAGTGTTCCTCGCCGAGACCTTACTGTGCTTGGGATGTTCT	1819
Db	538	GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAsn	557
Qy	1820	GCATGTTCTCGCTATTTCCACTGCAAGAGAGAGCAGCAGCACAAGATATAAGAAAT	1879
Db	558	AlaCysSerArgTyrAlaProThrSerLysArgAlaArgGlnAspValLysTyr	577
Qy	1880	GGAGACCATGACTCACTGCTTCCAGACTTACCATGATTAATCACCATGCCACACGCCCT	1939
Db	578	GlyAspProIleThrGlnCysTrpAspIleGluAspSerHis--GluThrAla	596
Qy	1940	GAAGAGAGAAATCATCTATGCTGTAGAGAAATAGTAGCACATTTTGGAAATGCAGTCCGAAG	1999
Db	597	AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys	616
Qy	2000	TCGCAGAGACCGCTGGTCTATTGGCAATTCAGAGCGGAAATGCAAGAGCGGAAAGAGAG	2059
Db	617	SerGlnGlnAlaThrIleLysTrpTyrIleGlnArgSerGlyAspGluHisArgGluGlu	636
Qy	2060	ATCAGAGTGATGATCATCATCATCAGCAGAGATCAAGGCTCTCTGCTAGCTAGTCTACAA	2119
Db	637	LeuLysProAspGluArgIleIleLysThrGluLysGlyLeuLeuIleArgSerLeuGln	656
Qy	2120	CAGAAGATTTCAGGCAATTACCTCTGCCATCGCGTGGAAACATGGGTTCATACAAACTCTT	2179
Db	657	LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle	676
Qy	2180	CTTAAGGTAAACCTTGAAGTCATTGACACAGACATTTGGAAGAACTTCTTCATAAGAT	2239
Db	677	ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu	696
Qy	2240	GATCATGGAGATGGCTCTAGACCAACAAAGAAATGTCCAATAGCATGACACCTAGCCAGAAG	2299
Db	697	HisGluGluGlyGlnValLys-----AspIleuAlaGluSerArg	710
Qy	2300	GTCTGGTACAGACATTCATCGAGCTCATCAACACCCCAATCTCAACAGCATGGATGAG	2359
Db	711	LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln	729
Qy	2360	TTCTGTCAACAAGTTTGGAAAAGGGACCGGAAAAACAACGTCGCGCAAGAGCCAGACATACC	2419
Db	730	TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly-----	747
Qy	2420	CCAGGAAACAGTAACAAATCGAGACCTTACAGAAATAAGAAAGGTGAACACGAGG	2479
Db	748	-----ProLysTrpLysHisMetGlnGluMetLysLysArgAsnArgArg	763
Qy	2480	ACCCAC---GAATTTGAGAGCGCACCCAGGAGTGC	2512
Db	764	HisHisArgAspLeuAspGluLeuProArgAlaVal	775

RESULT 8
RAU29197

Db 239 AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrThyrAsnProAspAsp 258
QY 923 AAAGTATACCTTTCTTCGCGTGAATGCAATAGATGGAGACACTCTGAAAGCTACT 982
Db 259 LysIleTyrPhePhePheArgGluSerGlnGlySerThrSerAspLysThrIle 278
QY 983 CAGCTTAGAATAGTTCAGATATGCAAGATGACTTTGGAGGCGCACAGAGTCTGCTGAAT 1042
Db 279 LeuSerArgValGlyArgValCysLysAsnAspValGlyGlnArgSerLeuIleAsn 298
QY 1043 AAATGACAACTTCCTCAAGCTCGTCTGATTGCTCAGTCCAGTCCCAATGCGCAT 1102
Db 299 LysTyrThrPheLeuLysAlaArgLeuIleCysSerIleProGlySerAspGlyAla 318
QY 1103 GACACTCATTTGATGAACCTGAGATGATTCCTTAATCAACTTAAAGATCCTAAAT 1162
Db 319 AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArgAsn 338
QY 1163 CCAGTTGTATATGGAGTGTTCAGACTTCAGCTTAACATTTTCAAGGATCAGCCGTGT 1222
Db 339 ProValValTyrGlyValPheThrThrSerSerIlePheLysGlySerAlaValCys 358
QY 1223 ATGTATAGCATGAGTGAATGAGAAGGTGTTCTTGTGTCATATGCCACAGGATGGA 1282
Db 359 ValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysGluSer 378
QY 1283 CCAACTATCATGGTGCCTTATCAGGAAGTCCCTATCCAGGAGCCAGGACTTGT 1342
Db 379 AlaAspHisArgTyrValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys 398
QY 1343 CCAGCAAAACATTTGGTGGT--TTTGACTCTCAAAAGGAGCTTCTGTATCATGTATA 1399
Db 399 ProSerLysThrTyrAspProLeuIleLysSerThrArgAspPheProAspValIle 418
QY 1400 ACCTTTGCAAGAGTCATCCAGCCATGTAATCAATCCAGTGTTCCTATGAACAAATGCCCA 1459
Db 419 SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro 438
QY 1460 ATAGTATCAAAACGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1519
Db 439 ThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIle 458
QY 1520 CGAAGATGAGCAGTATGATGTTATGTTTATCGGAACAGATGTTGGGACCGCTCTTAAA 1579
Db 459 AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
QY 1580 GTAGTTTCAATTCCTAAGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAAGAAATG 1639
Db 479 ValValSerIleSerLysGluLysTyr--AsnMetGluGluValValLeuGluGluLeu 497
QY 1640 ACAGTTTTTCGGGACCGACTGCTATTTCAGCAATGGAGCTTCCACTAAGCAGCAACAA 1699
Db 498 GlnIlePheLysHisSerSerIleLeuAsnMetGluLeuSerLysGlnGlnGln 517
QY 1700 CTATATATTGTTCAACGGCTGGGTTGCCAGCTCCCTTTTACCGGTGTGATATTAC 1759
Db 518 LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
QY 1760 GGAAGAGCGTGTGCTGAGTGTTCCTCGCCGAGACCCCTTACTGCTGGATGGTCT 1819
Db 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTyrAspGlyAsn 557
QY 1820 GCATGCTTCGCTATTTTCCCACTGCAAGAGACCCACAGACGACAGATATAGAAT 1879
Db 558 AlaCysSerArgTyrAlaProThrSerLysArgAlaArgArgGlnAspValLysTyr 577
QY 1880 GGAGACCCACTCACTCACTGTTTCAGACTTACACCATGATAATCACCAGCCAGCCCT 1939
Db 578 GlyAspProIleThrGlnCysTyrAspIleGluAspSerIleSerHis---GluThrAla 596
QY 1940 GAAGAGAGATCATCTATGCTGTAGAGATATAGACATTTTGGATGCTGAGTCCGAG 1999
Db 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616

QY 2000 TCGCAGAGAGCGCTGCTCTATTGGCAATTCCAGAGCGCAATGAAGACGCAAAAGAG 2059
Db 617 SerGlnGlnAlaThrIleLysTyrTyrIleGlnArgSerGlyAspGluHisArgGluGlu 636
QY 2060 ATCAGATGGATGATCATCATCATCAGGACAGATCAAGSCCTTCTGCTACGTAGTACAA 2119
Db 637 LeuLysProAspGluArgIleIleLysThrGlyTyrGlyLeuLeuIleArgSerLeuGln 656
QY 2120 CAGAAGATTTCAGCAATTACCTCTGCCATCGGTGGAAACATGGGTTTCATACAACTCTT 2179
Db 657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
QY 2180 CTTAAGGTAAACCTCGGAAGTCATGACACAGACATTTGGAGAACTTCTTCATAAAGAT 2239
Db 677 ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu 696
QY 2240 GATGATGAGATGGCTCTTAAGACCAAAAGAAATGTCATAGCATGACACCTAGCCAGAAG 2299
Db 697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
QY 2300 GTCTGGTACAGACTTCATCAGCTCATCAACACCCCAATCTCAACACAGATGATGAG 2359
Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln 729
QY 2360 TTCTGTGAACAAGTTTGGAAAGGACCGCAAAACAAACGTCGCAAGCCAGGACATACC 2419
Db 730 TyrCysGluGlnMetTyrPheArgGluLysArgGlnArgAsnLysGlyGly----- 747
QY 2420 CCAGGGAACATAACAATGGAAGCCTTACAAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 748 -----ProLysTyrLysHisMetGlnGluMetLysLysLysArgAsnArgArg 763
QY 2480 ACCAC---GAATTTGAGAGGCGCACCCAGGAGTGTCT 2512
Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775

RESULT 9

AA866176
ID AA866176 standard; protein; 777 AA.
AC AA866176;
XX
DT 02-APR-2001 (first entry)
XX
DE Protein of the invention #88.
XX
KW Secreted; transmembrane; gene therapy.
XX
OS Unidentified.
XX
PN WO20007961-A1.
XX
PD 28-DEC-2000.
XX
PF 18-FEB-2000; 2000WO-US04342.
XX
PR 23-JUN-1999; 99US-0141037.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 01-SEP-1999; 99WO-US20111.
PR 20-OCT-1999; 99US-0162506.
PR 30-NOV-1999; 99WO-US28313.
PR 02-DEC-1999; 99WO-US28551.
PR 16-DEC-1999; 99WO-US30095.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00376.
XX
XX (GETH) GENENTECH INC.

XX Baker KP, Botstein D, Desnoyers L, Eaton DL, Ferrara N, Fong S;
PI Gao W, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;
PI Pan J, Paoni NF, Roy MA, Smith V, Stewart TA, Tumas D;


```
QY 2000 TCGCAGAGAGCGCTGCTATTGGCAATTCACAGAGCGAAATGAAGAGCGAAAGAGAG 2059
Db 617 SerGlnGlnAlaThrIleLeuTyrTrpTyrIleGlnArgSerGlyAspGluHisArgGluGlu 636
QY 2060 ATCAGAGTGGATGATCATATCATCATCAGGACAGATCAAGGCTTCTCTAGCTAGTGTACAA 2119
Db 637 LeuLysProAspGluArgIleLeuLysThrGluTyrGlyLeuLeuLeuLeuLeuLeuLeu 656
QY 2120 CAGAGGATTCAGCAATTAATCTCTGCCATGCGGTGGACATGGGTTTATACAAATCTTT 2179
Db 657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
QY 2180 CTTAAGGTAACCTCGAGTCTATTCACAGAGCATTTGGAGAACTTCTTCATAAAGAT 2239
Db 677 ValLysLeuThrLeuAsnValIleGluAsnGluMetGluAsnThrGlnArgAlaGlu 696
QY 2240 GATGATGGAGATGGCTCTTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAGAAG 2299
Db 697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
QY 2300 GTCGTGTACAGAGCTTCATGAGCTCATCAACACCCCAATCTCAACAGATGATGAG 2359
Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln 729
QY 2360 TTCTGTGAACAAGTTTGGAAAGGACCGAAGCAAGCTCGGCAAGGCCAGGACATACC 2419
Db 730 TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly----- 747
QY 2420 CCAGGGAACAGTAAACAAATGGAAGCACCTTACAGAAATAAGAAAGGTAGAAACAGGAG 2479
Db 748 -----ProLysTrpLysHisMetGlnGluMetLysLysLysArgAsnArgArg 763
QY 2480 ACCCAC---GAATTTGAGAGGCCACCCAGGAGTGTC 2512
Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775

RESULT 10
ABU71285
ID ABU71285 standard; Protein; 777 AA.
XX
AC ABU71285;
DT 10-JUN-2003 (first entry)
XX
DE Human PRO1491 protein.
XX
KW Human; PRO; secreted; transmembrane; cytosolic; TNF-alpha; blood;
KW tumour necrosis factor alpha release; chondrocyte cell; proliferation;
KW differentiation; tumour; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003036143-A1.
XX
PD 20-FEB-2003.
XX
PF 02-JUL-2002; 2002US-0187600.
XX
PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15284.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 28-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 28-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
```


PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086021P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088325P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.

```

PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096881P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.

Alignment Scores:
Pred. No.: 2,688-198 Length: 777
Score: 2245.50 Matches: 413
Percent Similarity: 73.67% Conservative: 141
Best Local Similarity: 54.92% Mismatches: 177
Query Match: 45.86% Indels: 21
Db: 24 Gaps: 9

US-09-774-490-1 (1-2709) x ABU65742 (1-777)

QY 275 RAGAACATGTCAGAGGCTGAATATCTCTACAAAGAAATGTTGGAATCCAAACATGTG 334
Db 39 LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuLeuSerAsnSerCys 58
QY 335 ATCACTTTCAATGGCTGGCCCAACAGCTCCAGTTATCATACCTCTCTTTGGATGAGAA 394
Db 59 IleProLeuLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuLeuAspGlu 78
QY 395 CGAGTAGGCTGTATGTGGAGCAAGGATCATATTTTCATTCGACTGGTTAATATC 454
Db 79 ArgGlyArgLeuLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
QY 455 ---AAGGATTTTCAAAAGATCTGTGGCCAGTATCTTACACAGAGAGAGTGAATGCAAG 511
Db 99 AsnLysAsnPheLysIleTyrTrpProAlaLysGluArgValGluLeuCysLys 118
QY 512 TGGGCTGGAAGACATCTGAAAGATGTGCTAATTTTCATCAAGGTACTTAAGCATAT 571
Db 119 LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr 138
QY 572 ATCAAGACTCACTGTAGCCCTGTGGAGCGGGCTTTTCATCCAAATTTGCACCTACATT 631
Db 139 AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle 158
QY 632 GAAATTGGACATCATCTGAGGACAAATATTTTAAAGCTGGAGAACTCCATATTTGAAAC 691
Db 159 AspLeuGlyValTyrLysGluAspIlePheLysLeuAspThrHisAsnLeuGluSer 178
QY 692 GCCCGTGGGAAGAGTCCATATGACCTTAAGCTGTGCACAGCATCCCTTTTAAATAGATGA 751
Db 179 GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu 198
QY 752 GAATTATCTCTGGAAGTCCAGTGTATTTTATGGGCGAGACTTTTGTCTATCTTCCGAAT 811
Db 199 TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer 218
QY 812 CTTGGG-----CACCACCAACCAATCAGGACGACCATGATTTCCAGGTGGCTC 862
Db 219 LeuGlyProThrHisAspHisIleTyrIleArgThrAspIleSerGluHisTyrTrpLeu 238
QY 863 AATGATCCAAAGTTCAATAGTCCCACTTCATCTCAGAGAGTGACAACTCCTGAAGATGAC 922
Db 239 AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrAsnProAspAsp 258

```

```

QY 923 AAGTATATCTTTTCTTCCGTGAAATGCAATAGATGGAGAACACTCTGGAAGACTACT 982
Db 259 LysIleTyrPhePheArgGluSerGlnGlnGlySerThrSerAspLysThrIle 278
QY 983 CACGCTAGAATAGCTCAGATATGCAAGATGACTTTTGGAGGCCACAGAGTCTGTGTAAT 1042
Db 279 LeuSerArgValGlyArgValCysLysAsnAspValGlyGlyGlnArgSerLeuIleAsn 298
QY 1043 AATGGACAACATTCCTCAAGAGCTCGTCTGATTTTGTCTCAGTCCAGGTCCTCAATGGCATT 1102
Db 299 LysTrpThrPheLeuLysAlaArgLeuIleCysSerIleProGlySerAspGlyAla 318
QY 1103 GACACTCATTTTGAATGACGACGATGTATTCCTAATGAATCTTAAAGATCTTAAAGAT 1162
Db 319 AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArgAsn 338
QY 1163 CCAGTTGTATATGAGTGTTTTACGACTTCCAGTAAACATTTTCAAGGGATCAGCGGTGTGT 1222
Db 339 ProValValTyrGlyValPheThrThrSerSerIlePheLysGlySerAlaValCys 358
QY 1223 ATGTATAGCATGATGATGAGAGAGGTGTTCCTTGTCTCATATGCCACAGGATGGA 1282
Db 359 ValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysGluSer 378
QY 1283 CCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCAGGCCCAGGAACTTGT 1342
Db 379 AlaAspHisArgTrpValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys 398
QY 1343 CCCAGCAAAACATTTGGTGGT---TTTGACTCTCAAAAGAGCCTTCTCTGATGATGTTATA 1399
Db 399 ProSerLysThrTyrAspProLeuIleLysSerThrArgAspPheProAspValIle 418
QY 1400 ACCTTTGCAAGAAGTCTACGCCATGTACCAATCCAGTCTTCTTATCAACAATGCCCA 1459
Db 419 SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro 438
QY 1460 ATAGTGATCAAAACGGATGTAATATCAATTTTACACAAATTTGCTGAGACCGAGTGGAT 1519
Db 439 ThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIle 458
QY 1520 GCAGAAGATGACAGTATGATGTTATGTTATCGAAACAGATGTTGGAGCCGTTCTTAAA 1579
Db 459 AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAGAGGTTCTCTCGAAGAAATG 1639
Db 479 ValValSerIleSerLysGluLysTrp---AsnMetGluGluValValLeuGluGluLeu 497
QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTTCAGCAATGAGACTTTCACCTTAAGCAGCAACA 1699
Db 498 GlnIlePheLysHisSerSerIleIleLeuAsnMetGluLeuSerLeuLysGlnGlnGln 517
QY 1700 CTATATATGTTTCAACGGCTGGGTTGCCAGTCTCCCTTTTACACCGGTGTGATTTTAC 1759
Db 518 LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
QY 1760 GGGAAAGCGTGTCTGATGTTGCTCGCCGAGACCTTACTGCTGGATGGATGTTCT 1819
Db 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAsn 557
QY 1820 GCATGTTCTCGCTATTTTCCCACTGCAAGAGAGCGCACAGACGACAAAGATATAGAAT 1879
Db 558 AlaCysSerArgTyrAlaProThrSerLysArgArgAlaArgGlnAspValLysTyr 577
QY 1880 GGAGACCCACTGACTCACTGTTCAGACTTACACCATGATATACCATGGCCACAGCCCT 1939
Db 578 GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GluThrAla 596
QY 1940 GAAGAGAGAAATCATCTGTTGTTAGAGATAGTACCATTTTGGATGATCGAGTCCGAG 1999
Db 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616

```



```
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088036P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088555P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 10-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089808P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 26-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 04-JUL-1998; 98US-094006P.
PR 10-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.

PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096953P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 10-SEP-1998; 98US-099754P.
PR 10-SEP-1998; 98US-099763P.

Alignment Scores:
Pred. NO.: 2.68e-198
Score: 2245.50
Percent Similarity: 73.67%
Best Local Similarity: 54.92%
Query Match: 45.86%
DB: 24

US-09-774-490-1 (1-2709) x ABU66075 (1-777)

QY 275 AAGAACAAATGTCACAGGCTGAATAATCTCTACAAAGAAATGTTGGAATCCAAATGTG 334
DB 39 LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuSerAsnSerCys 58
QY 335 ATCATTTCATGCTGGCCACAGCTCCAGTTATCATACCTTCCTTTTGGATGAGGAA 394
DB 59 IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuAspGluGlu 78
QY 395 CGGAGTAGGCTGATGTTGGAGCAAGATCACAATATTTTCATTCGACCTGGTTAATATC 454
DB 79 ArgGlyArgLeuLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
QY 455 ---AAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACACCAAGAGAGATGAATGCAAG 511
DB 99 AsnLysAsnPheLysLysIleTyrTyrProAlaAlaLysGluArgValGluLeuCysLys 118
QY 512 TGGCTCGAAAGACATCCTGAAAGATGTCTAATTTTCATCAAGTACTTTAAGGCATAT 571
DB 119 LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr 138
QY 572 AATCAGACTCATTGTAGCGCTGTGGACGGGGCTTTTCATCCCAATTTGCACCTACATT 631
DB 139 AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle 158
QY 632 GAAATTCGACATCATCTCTAGGACAAATATTTTAAAGCTGGAGAACTCACATTTTCAAAAC 691
DB 159 AspLeuGlyValTyrLysGluAspIleIlePheLysLeuAspThrHisAsnLeuGluSer 178
QY 692 GGCGTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAAATAGATGA 751
DB 179 GlyArgLeuLysCysProPheAspProGlnProPheAlaSerValMetThrAspGlu 198
QY 752 GAATTATCTCTGGAAGTGGATTTTATGGGCGAGACTTTGTGCTATCTTCGAGCT 811
DB 199 TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer 218
QY 812 CTTGGG-----CACCACCAACCCATCAGACAGACAGCAGCATGATTCAGGTGGCTC 862
DB 219 LeuGlyProThrHisAspHisHisTyrIleArgThrAspIleSerGluHisTyrTrpLeu 238
QY 863 AATGATCCAAAGTTTCATTAGTGGCCCACTCATCTCAGAGAGTGACATCTCTGAGATGAC 922
DB 239 AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrAsnProAspAspAsp 258
```

Qy	2000	TCGCAGAGACCGCTGGTGTCTATTTCGCCAATTCACAGAGCGCAAAATGAAAGACGCGAAAGAGAG	2059
Db	617	SerGlnGlnAlaThrIleLysTyrPtyrIleGlnArgSerGlyAspGluHisArgGluGlu	636
Qy	2060	ATCAGAGTGGATGATCATATCATCAGGACAGATCAAGCGCTCTCTGCTAGCTAGCTCTACAA	2119
Db	637	LeuLysProAspGluArgIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln	656
Qy	2120	CAGAAGGATTTCAGGCAATTAACCTCTCGCATGCGGTGGAACATGGGTTCATACAACTCTT	2179
Db	657	LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle	676
Qy	2180	CTTAAGGTAAACCTCGAAGTCATGTACACAGACATTTGGAAGAACTTCTTCAATAAGAT	2239
Db	677	ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu	696
Qy	2240	GATGATGGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAGAG	2299
Db	697	HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg	710
Qy	2300	GTCTGGTACAGACATTCATGACAGCTCATCAACACCCCAATCTCAACACGATGGATGAG	2359
Db	711	LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln	729
Qy	2360	TTCTGTGAACAAGTTTGGAAAAGGGACCGAAACAACGTCGCGCAAGGCCAGGACATACC	2419
Db	730	TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly-----	747
Qy	2420	CCAGGGACAGTAACAATGGAAGCACTTACAGAATAAGAAAGGTAGAAACAGCAGG	2479
Db	748	-----ProLysTrpLysHisMetGlnGluMetLysLysLysArgAsnArgArg	763
Qy	2480	ACCCAC---GAATTTTCGAGAGGCGCCACGAGGTGTC	2512
Db	764	HisHisArgAspLeuAspGluLeuProArgAlaVal	775
RESULT 13			
ABU67579			
ID	ABU67579	standard; Protein; 777 AA.	
XX	AC	ABU67579;	
XX	XX		
DT	29-MAY-2003	(first entry)	
XX			
DE		Human secreted/transmembrane protein (PRO) #174.	
XX			
KW		Human; secreted and transmembrane protein; PRO; TNF-alpha;	
KW		tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;	
KW		tissue typing.	
XX			
OS		Homo sapiens.	
XX			
PN	US2003036162-A1.		
XX			
PD	20-FEB-2003.		
XX			
PF	12-JUL-2002; 2002US-0194423.		
XX			
PR	16-SEP-1998;	98WO-US19330.	
PR	07-OCT-1998;	98WO-US21141.	
PR	01-DEC-1998;	98WO-US25108.	
PR	08-MAR-1999;	99WO-US05028.	
PR	14-MAY-1999;	99WO-US10733.	
PR	02-JUN-1999;	99WO-US12252.	
PR	01-SEP-1999;	99WO-US20111.	
PR	15-SEP-1999;	99WO-US21090.	
PR	01-DEC-1999;	99WO-US28301.	
PR	02-DEC-1999;	99WO-US28551.	
PR	30-DEC-1999;	99WO-US31274.	
PR	05-JAN-2000;	2000WO-US00219.	
PR	18-FEB-2000;	2000WO-US04341.	
PR	18-FEB-2000;	2000WO-US04342.	
PR	22-FEB-2000;	2000WO-US04414.	

24-FEB-2000; 2000WO-US05004.
 01-MAR-2000; 2000WO-US05601.
 02-MAR-2000; 2000WO-US05641.
 15-MAR-2000; 2000WO-US05684.
 30-MAR-2000; 2000WO-US08439.
 17-MAY-2000; 2000WO-US13705.
 22-MAY-2000; 2000WO-US14042.
 30-MAY-2000; 2000WO-US14941.
 02-JUN-2000; 2000WO-US15264.
 28-JUL-2000; 2000WO-US20710.
 24-AUG-2000; 2000WO-US23328.
 08-NOV-2000; 2000WO-US30952.
 01-DEC-2000; 2000WO-US32678.
 20-DEC-2000; 2000WO-US34956.
 28-FEB-2001; 2001WO-US05520.
 01-JUN-2001; 2001WO-US17800.
 20-JUN-2001; 2001WO-US19692.
 29-JUN-2001; 2001WO-US21066.
 09-JUL-2001; 2001WO-US21735.
 29-AUG-2001; 2001WO-US27099.
 26-JUN-1998; 98US-0105413.
 07-OCT-1998; 98US-0168978.
 06-NOV-1998; 98US-0187368.
 07-DEC-1998; 98US-0202054.
 03-MAR-1999; 98US-0254311.
 14-MAY-1999; 98US-0311832.
 14-MAY-1999; 98US-0380137.
 25-AUG-1999; 98US-0380138.
 25-AUG-1999; 98US-0380139.
 25-AUG-1999; 98US-0380142.
 18-OCT-1999; 98US-0403297.
 12-NOV-1999; 98US-0423844.
 22-AUG-2000; 2000US-0644848.
 18-SEP-2000; 2000US-0664610.
 18-SEP-2000; 2000US-0665350.
 08-NOV-2000; 2000US-0709238.
 20-DEC-2000; 2000US-0747259.
 22-MAR-2001; 2001US-0816744.
 10-MAY-2001; 2001US-0854208.
 10-MAY-2001; 2001US-0854280.
 25-MAY-2001; 2001US-0866028.
 05-JUN-2001; 2001US-0874503.
 18-JUL-2001; 2001US-0908827.
 30-JUL-2001; 2001US-0918585.
 06-AUG-2001; 2001US-0924419.
 13-AUG-2001; 2001US-0929404.
 16-AUG-2001; 2001US-0931836.
 28-AUG-2001; 2001US-0941992.
 04-SEP-2001; 2001US-0946374.
 15-JAN-2002; 2002US-0052586.
 (GETH) GENENTECH INC.
 Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
 Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
 WPI; 2003-332039/31.
 N-PSDB; ACA05873.
 New secreted and transmembrane PRO polypeptides and nucleic acids,
 useful in gene therapy, in chromosome and gene mapping, as chromosome
 markers, in tissue typing, and in chromosome identification -
 Claim 11; Fig 348; 706pp; English.
 The invention discloses human nucleic acids encoding secreted and
 transmembrane (PRO) polypeptides. Also disclosed is an antibody that
 specifically binds to the PRO polypeptide, a method for stimulating the
 release of tumour necrosis factor alpha (TNF-alpha) from human blood by
 contacting the blood a PRO polypeptide, a method for stimulating the
 proliferation or differentiation of chondrocyte cells by contacting the
 cells with a PRO polypeptide, a method for detecting the presence of a
 tumour in a mammal and an oligonucleotide probe derived from any of the

CC PRO nucleotide sequences. The nucleotide sequences are useful as probes,
 CC in chromosome and gene mapping, in generating antisense RNA and DNA, in
 CC preparing PRO polypeptides by recombinant techniques and in gene therapy
 CC (e.g. for replacement of defective gene). The PRO polypeptides are useful
 CC as molecular weight markers for protein electrophoresis purposes, for
 CC chromosome identification, as chromosome markers, as therapeutic agents,
 CC for stimulating the release of TNF-alpha from human blood, for
 CC stimulating the proliferation or differentiation of chondrocytes and
 CC detecting the presence of a tumour. The PRO polypeptides and nucleic
 CC acids may also be used diagnostically for tissue typing. The sequences
 CC presented in ABU67406-ABU67710 are the PRO polypeptides of the invention.
 XX

SQ Sequence 777 AA;

Alignment Scores:
 Pred. No.: 2,68e-198 Length: 777
 Score: 2245.50 Matches: 413
 Percent Similarity: 73.67% Conservative: 141
 Best Local Similarity: 54.92% Mismatches: 177
 Query Match: 45.86% Indels: 21
 DB: 24 Gaps: 9

US-09-774-490-1 (1-2709) x ABU67579 (1-777)

QY 275 AAGAACAAATGTCACAGGCTGAAATATCTCTACAAAGAAATGTTGGAAATCCAAACAATGTG 334
 DB 39 LysGlnAsnIleProArgLeuLysLeuThrTyrlsAspLeuLeuLeuSerAsnSerCys 58
 QY 335 ATCACTTTCAATGCTTGGCCACAGCTCCAGTTATCATCATCTCTCTCTTTTGATGAGAA 394
 DB 59 IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuLeuAspGlu 78
 QY 395 CGGAGTAGGCTGATGTTGGAGCAAGATCAGATATTTTCATTCGACCTGGTTAATATC 454
 DB 79 ArgGlyArgLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
 QY 455 ---AAGGATTTCAAAGATTTGTGTGCCATCTTACACGAGAGAGATGAATGCAAG 511
 DB 99 AsnLysAsnPheLysIleTyrlProAlaAlaLysGluArgValGluLeuLys 118
 QY 512 TGGCTGGAAAGACATCTCTGAAAGATGTCTAATTTCAATCAAGGTACTTTAAGGCATAT 571
 DB 119 LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr 138
 QY 572 AATCAGACTCTGTGTAGCCTGTGGACGGGGCTTTTCATCCAAATTTGACCTACATTT 631
 DB 139 AsnLysThrHisIleTyrlValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle 158
 QY 632 GAAATTGGACATCATCTGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTTGAAC 691
 DB 159 AspLeuGlyValTyrlLysGluAspIleIlePheLysLeuAspThrHisAsnLeuGluSer 178
 QY 692 GGCCGTGGAGAGTCCATATGACCTTAAGCTGTGACAGCATCCCTTTTAAATAGATGGA 751
 DB 179 GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu 198
 QY 752 GAATTATCTCTGAACTCGAGCTGATTTATGGGCGAGACTTTGCTATCTTCGAACT 811
 DB 199 TyrLeuTyrlSerGlyThrAlaSerAspPheLeuGlyLysAspThrThrAlaPheThrArgSer 218
 QY 812 CTTGGG-----CACCACCCCAATCAGGACAGAGCAGCATGATTCCAGGTGGCTC 862
 DB 219 LeuGlyProThrHisAspHisIleTyrlArgThrAspIleSerGluHisTyrlTrpLeu 238
 QY 863 AATGATCAAAAGTTTCATTAGTGGCCCATCTCATCGAGAGTGACATCTCGAAGATGAC 922
 DB 239 AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrlAsnProAspAsp 258
 QY 923 AAAGTATCTTTTCTCCGTGAAATGCAATAGATGAGACACTCTGGAAAGCTACT 982
 DB 259 LysIleTyrlPhePheArgGluSerSerGlnGluGlySerThrSerAspLysThrIle 278
 QY 983 CACGCTAGAATAGTGTGATATGCAAGATGACTTTTGGAGGGCACAGAACTCTGTGTAAT 1042

Db	637	LeuLysProAspGluArgllelleLysThrGluTyrGlyLeuLeulleArgSerLeuGln	656
Qy	2120	CAGAAGGATTTCAGGCAATTACCTCTGCCATGCGGTGGAAACATGGGTTCATACAAACTCTT	2179
Db	657	LyLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle	676
Qy	2180	CTTAAGGTAAACCTGGAAGTCATTGACACAGACGACATTTGGAAGAACCTCTTCTATAAAGAT	2239
Db	677	VallYsLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu	696
Qy	2240	GATGATGGAGATGGCTCTAAGACCAAGAAATGTCTCAATAGCATGACACCTAGCCAGAAAG	2299
Db	697	HisGluGluGlyClnVallys-----AspIeuLeuAlaGluSerArg	710
Qy	2300	GTCTGTTACAGAGACTTCATGCACTCATCAACACCCCAATCTCAACACGATGGATGAG	2359
Db	711	LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerIeuAspGln	729
Qy	2360	TTCTGTGAACAAGTTTGGAAAGGGACCGAAACAACGTCGGCAAAAGCCAGGACATACC	2419
Db	730	TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly-----	747
Qy	2420	CCAGGGACACAGTAACAAATGGAGCACTTACAGAAATAAGAAAGGTAGAAACAGGAGG	2479
Db	748	-----ProLysTrpLysHisMetGlnGluMetLysLysArgAsnA-gArg	763
Qy	2480	ACCCAC---GAATTTTGAGAGGGCACCCAGGAGTGTC	2512
Db	764	HisHisArgAspLeuAspGluLeuProArgAlaVal	775
RESULT 14			
ABU65437			
ABU65437		standard; Protein; 777 AA.	
XX	AC	ABU65437;	
XX	AC		
DT	16-MAY-2003	(first entry)	
DE	Human PRO	polypeptide #174.	
XX	Human; PRO;	cytostatic; chromosome mapping; gene mapping;	
KW	protein electrophoresis;	tumour necrosis factor-alpha; TNF-alpha; blood;	
KW	chondrocyte differentiation;	chondrocyte proliferation; tumour.	
XX	Homo sapiens.		
OS	US2003032102-A1.		
PN	XX		
PF	13-FEB-2003.		
XX	17-JUN-2002;	2002US-0173697.	
XX	16-SEP-1998;	98WO-US19330.	
PR	07-OCT-1998;	98WO-US21141.	
PR	01-DEC-1998;	98WO-US25108.	
PR	08-MAR-1999;	99WO-US05028.	
PR	14-MAY-1999;	99WO-US10733.	
PR	02-JUN-1999;	99WO-US12252.	
PR	01-SEP-1999;	99WO-US20111.	
PR	15-SEP-1999;	99WO-US21090.	
PR	01-DEC-1999;	99WO-US28301.	
PR	02-DEC-1999;	99WO-US28551.	
PR	30-DEC-1999;	99WO-US31274.	
PR	05-JAN-2000;	2000WO-US00219.	
PR	18-FEB-2000;	2000WO-US04341.	
PR	19-FEB-2000;	2000WO-US04342.	
PR	22-FEB-2000;	2000WO-US04414.	
PR	24-FEB-2000;	2000WO-US05004.	
PR	01-MAR-2000;	2000WO-US05601.	
PR	02-MAR-2000;	2000WO-US05841.	
PR	15-MAR-2000;	2000WO-US06884.	
PR	30-MAR-2000;	2000WO-US08439.	
PR	17-MAY-2000;	2000WO-US13705.	

PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US22328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US05520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059466P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063540P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078986P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088126P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 02-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096894P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.

Db	637	LeuLysProAspGluAArgIleLeLysThrGluTyrGlyLeuLeuIleAArgSerLeuGln	656
Qy	2120	CAGAAGGATTCAAGCAATTACTCTGCATGCGGTGGAACATGGTTTCATACAAACCTCTT	2179
Db	657	LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle	676
Qy	2180	CTTAAGGTAACCCCTGGAGTCATTGACACAGACGATTTGGGAAGAACTTCTTCATAAAGAT	2239
Db	677	ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu	696
Qy	2240	GATGATGGAGATGGCTCTAAGACCAAGAAMAATGTCCAATAGCATGACACCTAGCCAGAAG	2299
Db	697	HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg	710
Qy	2300	GTCTGGTACAGAGACTTCATCGACTCATCAACCAACCCCAATCTCAACACGATGATGAG	2359
Db	711	LeuArgTyrLysAspTyrIleGlnIleuSerSerProAsnPhe---SerLeuAspGln	729
Qy	2360	TTCTGTGAACAAAGTTTGGAAAGGACCGCAAAACAACGTCGGCAAGGCCAGGACATACC	2419
Db	730	TyrCysGluGlnMetTyrHisArgLysArgArgGlnArgAsnLysGlyGly-----	747
Qy	2420	CCAGGGAACAGTACAAATGGAACACTTACAGAAAAATAGAAAAGGTAGAAACAGGAGG	2479
Db	748	-----ProLysTrpLysHisMetGlnGluMetLysLysLysArgAsnArgArg	763
Qy	2480	ACCCAC--GAATTTGAGAGGGCACCCAGGAGTGTC	2512
Db	764	HisHisArgAspLeuAspGluLeuProArgAlaVal	775
RESULT 15			
ABUS8573			
ID	ABUS8573	standard; Protein; 777 AA.	
AC	ABUS8573;		
DT	15-APR-2003	(first entry)	
DE	Human PRO polypeptide #174.		
KW	Human; PRO; cytostatic; tumour; cancer; breast; lung; stomach;		
KW	liver; dog; cat; cow; horse; sheep; pig; goat; rabbit; ADEPT;		
KW	antibody-dependent enzyme mediated prodrug therapy.		
OS	Homo sapiens.		
FN	US2003027272-A1.		
PD	06-FEB-2003.		
PF	21-JUN-2002; 2002US-0176492.		
XX	16-SEP-1998; 98WO-US19330.		
PR	07-OCT-1998; 98WO-US21141.		
PR	01-DEC-1998; 98WO-US25108.		
PR	08-MAR-1999; 99WO-US05028.		
PR	10-MAR-1999; 99WO-US05190.		
PR	14-MAY-1999; 99WO-US10733.		
PR	02-JUN-1999; 99WO-US12252.		
PR	01-SEP-1999; 99WO-US20111.		
PR	15-SEP-1999; 99WO-US21090.		
PR	30-NOV-1999; 99WO-US28313.		
PR	01-DEC-1999; 99WO-US28301.		
PR	02-DEC-1999; 99WO-US28551.		
PR	30-DEC-1999; 99WO-US31274.		
PR	05-JAN-2000; 2000WO-US00219.		
PR	18-FEB-2000; 2000WO-US04341.		
PR	18-FEB-2000; 2000WO-US04342.		
PR	22-FEB-2000; 2000WO-US04414.		
PR	24-FEB-2000; 2000WO-US05004.		
PR	01-MAR-2000; 2000WO-US05061.		
PR	02-MAR-2000; 2000WO-US05841.		
PR	10-MAR-2000; 2000WO-US06319.		


```
Db 279 LeuSerArgValGlyArgValCysLysAsnAspValGlyGlnArgSerLeuLeuAsn 298
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1043 AAATGGACAACATTCCTCAAGCTCTCTGATTTGCTAGTCCAGGTCCAATGECATT 1102
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 299 LysTrpThrPheLeuLysAlaArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 318
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1103 GACACTCATTTGATGAACCTGCAGGATGTATTCCTTAATGAACCTTAAAGATCCTAAAT 1162
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 319 AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArgAsn 338
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1163 CCAAGTTGATATGAGCTGTTTACGACTTCCAGTAACATTTTCAAGGATCAGCGGTGT 1222
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 339 ProValValTyrGlyValPheThrThrSerSerSerIlePheLysGlySerAlaValCys 358
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1223 ATGTATPAGCATGATGATGAGAGGGTGTCTTCTTGGTCCATATGCCACAGGATGGA 1282
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 359 ValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysGluSer 378
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1283 CCCAACTATCAATGGTGGCTTATCAAGGAGAGTCCCTATCCAGCGCCAGCACTTGT 1342
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 379 AlaAspHisArgTrpValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys 398
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1343 CCAGCAAAACATTTTGGTGGT---TTTGACTCTCAAAAGCAGCTTCTCATGATGTATA 1399
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 399 ProSerLysThrTyrAspProLeuLeuLysSerThrArgAspPheProAspValIle 418
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1400 ACTTTGCAAGATGATCCAGCCATGTAACATCCAGTGTTCCTATGAACCAATGCCCA 1459
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 419 SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro 438
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1460 ATAGTCATCAAAACGATGTAATATATCAATTTACACAATTTGCTAGACCGAGTGGAT 1519
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 439 ThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIle 458
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1520 GCAGAGATGGACAGTATGATGTTATGTTTATCGGAACAGATGTTGGGACCGTCTTAA 1579
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 459 AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1580 GTATTTTCAATTCCTTAAGAGACTTGGTATGATTTAGAGAGGTCTGCTGGAAGAAATG 1639
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 479 ValValSerIleSerLysGluLysTrp---AsnMetGluGluValValLeuGluLeu 497
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTTCCACTAAGCAGCAACAA 1699
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 498 GlnIlePheLysHisSerSerIleLeuLeuAsnMetGluLeuSerLeuLysGlnGln 517
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1700 CTATATATTGGTTCAACGGTGGGGTTCACAGCTCCCTTTACACGGGTGATATTAT 1759
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 518 LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1760 GGAAGAGCTGTGCTGAGTGTTCCTCGCCGAGACCTTTACTGTCTTGGATGCTTCT 1819
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAsn 557
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1820 GCATGTTCTCGCTATTTTCCACTGCAAGAGACGCGACAGCAGCAGATATAGAAT 1879
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 558 AlaCysSerArgTyrAlaProThrSerLysArgAlaArgAlaArgGlnAspValLysTyr 577
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1880 GGAGACCCCACTGACTCACTGTTCCAGATTTACACCATGATAATCACCATGGCCACAGCCCT 1939
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 578 GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GluThrAla 596
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 1940 GAAGAGAGATCATCTATCGTGTAGAGATAGTAGCACAATTTTGAATGCGAGTCCGAG 1999
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 2000 TCCGAGAGCGCTGCTTATGGCAATTCGAGCGCGAAATGAGAGCGAAAGAGAG 2059
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 617 SerGlnGlnAlaThrIleLysTrpTyrIleGlnArgSerGlyAspGluHisArgGluGlu 636
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 2060 ATCAGAGTGGATCATCATCATCAGGACAGATCAAGCGCTTCTGCTAGTGTCTACAA 2119
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
```

Search completed: August 3, 2003, 10:05:38
Job time : 127.5 secs

```
Db 637 LeuLysProAspGluArgIleIleLysThrGluTyrGlyLeuLeuLeuArgSerLeuGln 656
QY 2120 CAGAGAGATTTCAGCAATTACTCTGCCATCGGTTGGAAACATGGGTTTCATACAACTCTT 2179
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 2180 CTTAAGGTAAACCTCGGAAGTCATTGACACAGCAGCATTTGGAAGAACTTCTTCATAAAGAT 2239
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 677 ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu 696
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 2240 GATGATGGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACACCTAGCCAGAG 2299
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 2300 GTCTGGTACAGAGACTTCATGCTCATCAACACCCCAATCTCAACACGATGGATGAG 2359
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerProAsnPhe---SerLeuAspGln 729
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 2360 TTCTGTGAACAAAGTTTGGAAAAGGACCGAAACAAACGTCGGCAAGGCCAGGACATACC 2419
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 730 TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly----- 747
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 2420 CCAGGGAACAGTAACAATGGAAGCATTACAGAAATAAGAAAGGTAGAAACAGAGAG 2479
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 748 -----ProLysTrpLysHisMetGlnGluMetLysLysLysArgAsnArgArg 763
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 2480 ACCCAC---GAATTTGAGAGGCGACCCAGGAGTCTC 2512
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 09:54:25 ; Search time 51 Seconds
(without alignments)

10216.508 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 4896

Sequence: 1 aatctttatttcatgatg.....aggcttttttctcctaataacc 2709

Scoring table:

BLAST62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 566616

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09774490/runat 03082003.095416.25340/app.query.fasta_1.2887
-DB=PIR 76 -START=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -QFMT=1 -END=1 -MATRIX=blowsum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09774490 @CNC 1.1.69 @runat 03082003.095416.25340 -NCPU=6 -ICPU=3
-NO.MWAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPOP=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4201	85.8	771	2	D49423
2	3989.5	81.5	772	2	I48747
3	3784.5	77.3	772	2	A49069
4	3478.5	71.0	666	2	I58169
5	2118	43.3	749	2	G01856
6	2059.5	42.1	748	2	I48744
7	1952	39.9	753	2	G02173
8	1850.5	37.8	751	2	I48748
9	966	19.7	834	2	S66498
10	857.5	17.5	782	2	I48746
11	802	16.4	760	2	I48745
12	785.5	16.0	724	2	C49423
13	766.5	15.7	730	2	JH0798
14	731.5	14.9	656	2	B49423

15	730	14.9	1074	2	JC5928	semaphorin F precu
16	699.5	14.3	712	2	T27165	hypothetical prote
17	684.5	14.0	711	2	A49423	semaphorin I precu
18	534	10.9	653	2	T03102	semaphorin homolog
19	327	6.7	676	2	T33853	hypothetical prote
20	197	4.0	1894	2	JC4980	plexin 1 precursor
21	188.502	3.9	1945	2	T13937	plexin A - fruit f
22	186	3.8	1872	2	JC4976	plexin 3 precursor
23	184	3.8	441	2	S29921	hypothetical prote
24	180	3.7	403	2	E42521	A39R protein - vac
25	172	3.5	1884	2	JC4975	plexin 2 precursor
26	170	3.5	1905	2	T51553	plexin - African c
27	167.002	3.4	2051	2	T13164	plexin B - fruit f
28	154.5	3.2	1568	2	T09074	semaphorin recepto
29	134.5	2.7	1806	2	T23298	hypothetical prote
30	127.5	2.6	295	2	J01775	Salt9R protein - v
31	123	2.5	573	1	CSBYT	catalase (EC 1.11.
32	120	2.5	904	2	S53856	DNA mismatch repai
33	119	2.4	1104	1	A36866	microbial collagen
34	118.5	2.4	1291	2	T09273	probable tail-host
35	118.5	2.4	1654	2	A12067	two-component sens
36	118	2.4	1379	1	S01254	hepatocyte growth
37	117.5	2.4	1029	2	H86179	hypothetical prote
38	115	2.3	3381	2	T42389	versican precursor
39	114.502	2.3	993	2	A36873	protein-tyrosine k
40	114	2.3	692	1	S57592	probable phosphoe
41	113	2.3	1276	2	T09204	probable tail-host
42	112.5	2.3	248	2	A46652	glucosamine-6-phos
43	112	2.3	4273	2	C59679	polyketide synthas
44	111.5	2.3	6642	2	T29757	protein UNC-89 - C
45	111	2.3	1375	2	T30813	plasmaonogen relate

ALIGNMENTS

RESULT 1

D49423

semaphorin III precursor - human

C:Species: Homo sapiens (man)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-Sep-1999

C:Accession: D49423

R:Kotodkin, A.B.; Matthes, D.J.; Goodman, C.S.

Cell 75, 1389-1399, 1993

A:Title: The Semaphorin genes encode a family of transmembrane and secreted growth cone

A:Reference number: A49423; MUID:94094332; PMID:8269517

A:Accession: D49423

A>Status: preliminary; nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-771 <KOL>

A:Cross-references: GB:L26081; NID:G799328; PIDN:AAA65938.1; PID:G436560

C:Genetics:

A:Gene: GDB:SEMA1

A:Cross-references: GDB:283448

C:Superfamily: semaphorin

Alignment Scores:

Pred. No.:	1.59e-310	Length:	771
Score:	4201.00	Matches:	771
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	85.80%	Indels:	0
DB:	2	Gaps:	0

US-09-774-490-1 (1-2709) x D49423 (1-771)

QY	200	ATGGCGTGGTAACTAGAGATTGCTCTTTCTGGGAGATTACTTACAGCAAGACA	259
Db	1	MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla	20
QY	260	AACATCAGATGGAGAACATGTCACAGGCTGAATTTATCTCTACAAAGAAATGTTG	319
Db	21	AsnTyrGlnAsnGlyLysAsnValProArgLeuLysLeuSeryrLysGluMetLeu	40

QY 320 GAATCCAAATGATGATCACTTTCAATGGCTGGCCAAACAGCTCCAGTTATCATACCTTC 379
Db |||||
41 GluSerAsnValIleThrPheAsnGlyLeuAlaAsnSerSerThrHisThrPhe 60
QY 380 CTTTTCGATGAGGAACGGAGTAGCTGTATGTTGGAGCAAGGATCACATATTTTCATTC 439
Db |||||
61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAATATCAAGATTTTCAAAAGATTGTGTGGCCAGTATCTTACACCGAAGA 499
Db |||||
81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATGAATCAAGTGGCTCGAAAGACATCTCCAAAGATGCTTAATTTTCATCAAGGTA 559
Db |||||
101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCACTTGTACGGCTGTGGAAACGGGGCTTTTCATCAAT 619
Db |||||
121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTGAATATGACATCATCTCTGAGGACAATATTTTAAAGCTGAGAACTCA 679
Db |||||
141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAACGGCGTGGAGAGTCCATATGACCTTAAGCTCTGACAGCATCCCTT 739
Db |||||
161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGGAATATTACTCTCGAACTGCGAGCTGATTTTATGGGGGAGACTTTGCT 799
Db |||||
181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTCCGAACTCTGGGACACACCAATCAGACAGAGAGAGATGATTCAGAGTGG 859
Db |||||
201 IlePheArgThrLeuGlyHisHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCATGATGCCAAGTTCTATTAGTGCACCTCATCTCAGAGAGTGCACATCTCAAGAT 919
Db |||||
221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATCTTTTCTCGTGAAATGCAATAGATGGAGAACACTCTCGAAAGCT 979
Db |||||
241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGATAGTATGATATCAAGATGATCTTGGAGGGACAGAACTCTGGTG 1039
Db |||||
261 ThrHisAlaArgIleGlyGlnLysCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAATGACACATCTCAAGCTCGTCTGATTTGCTCAGTGCAGGTCCCAATGGC 1099
Db |||||
281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTCATGAACATGAGGATGTATCTTAATGAACCTTTAAAGATCCTAAA 1159
Db |||||
301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 ATCCAGTGTATATGAGTGTATACGACTTCAGTAAACATTTTCAAGGGATCAGCCGTG 1219
Db |||||
321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGATGTGAGAAGGGTGTCTTGGTCTCATATGCCACAGGGAT 1279
Db |||||
341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCCACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTTATCCAGGGCAGGAAT 1339
Db |||||
361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValIleProTyrProArgProGlyThr 380
QY 1340 TGTCCAGCAAAACATTTGGTGGTTTGTACTTACAAAGAGACCTTCTGATGATGTTATA 1399
Db |||||
381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAAGAAGTATCCAGCCATGTACAATCCAGTGTCTTCTATGAACAATCGCCCA 1459

Db |||||
401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACGGATGTAATATCAATTTTACAAATTTGCTGTAGACCGAGTGGAT 1519
Db |||||
421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGNAGTGCACAGTATGATGTTATGTTATCGGAACAGATGTTGGACCGTCTTTAAA 1579
Db |||||
441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATCTCAAGGAGACTTGGTATGATTTAGAGAGGTTCTCTGTGAAGAAATG 1639
Db |||||
461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluMet 480
QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTCAGCAATGGAGCTTTCACTAAGCAGCAACAA 1699
Db |||||
481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATTGTTCAACGGCTGGGTTGCCAGCTCCCTTTACACCGGTGTGATTTTAC 1759
Db |||||
501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGCGTGTGCTGAGTGTGCTCGCCGAGACCCCTTACTGTGCTTGGGATGGTCT 1819
Db |||||
521 GlyLysAlaCysAlaGluCysLysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCACTGCAAGAGACGACACAGACGCAAGATATAAGAAAT 1879
Db |||||
541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgLysAspIleArgAsn 560
QY 1880 CGAGNCCCACTGACTCTGCTGAGTTCAGCTTACACATGATATCACCATGGCCACAGCCCT 1939
Db |||||
561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580
QY 1940 GAAGAGAAATCATCTATGTTGTAGAGAAATAGTAGACATTTTGGAAATGAGTCCGGAAG 1999
Db |||||
581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGAGCGCTGCTCTATTGGCAATTCAGAGCGGCAATGAAGAGCGGAAAGAGAG 2059
Db |||||
601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCCCTTCTGCTAGTAGTCTACAA 2119
Db |||||
621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGGATTCAGGCAATTTACTCTCTGCCATCGGTGGAAACATGGGTTTCATACAACTCTT 2179
Db |||||
641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGTAAACCTCGGAAGTCAATGACACAGAGCAATTTGGAGAACTTCTTCATAAAGAT 2239
Db |||||
661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuHisLysAsp 680
QY 2240 GATGATGAGATGGCTCTTAAGACCAAGAAATGCCATAGCATGACACCTAGCCAGAGAG 2299
Db |||||
681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGGTACAGAGACTTTCATGCGCTCATCAACCCCACTCTCAACACGATGGATGAG 2359
Db |||||
701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAAAGTTTGGAAAGGACCGGAAAACAACTCGGCCAAAGCCAGGACATACC 2419
Db |||||
721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGACACTTAACAATGAGACACTTACAAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db |||||
741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCCACCAATTTGAGAGGGCCACCCAGGAGTGTCT 2512
Db |||||

Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 2
I48747
semaphorin D - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C:Accession: I48747
R:Puschel, A.W.; Adams, R.H.; Betz, H.
Neuron 14, 941-948, 1995
A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates
A:Reference number: I48744; MUID:95267431; PMID:7748561
A:Accession: I48747
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-772 <RES>
A:Cross-references: EMBL:X85993; NID:g854329; PID:CAA59985.1; PID:g854330
C:Genetics:
C:Gene: semD
C:Superfamily: semaphorin

Alignment Scores:
Pred. No.: 1,76e-294 Length: 772
Score: 3989.50 Matches: 730
Percent Similarity: 97.28% Conservative: 21
Best Local Similarity: 94.56% Mismatches: 20
Query Match: 81.48% Indels: 1
DB: 2 Gaps: 1

US-09-774-490-1 (1-2709) x I48747 (1-772)

QY 200 ATGGCTGGTTAACTAGGATTGTCTCTTTCTGGGAGTATTACTTACAGCAAGACA 259
Db 1 MetGlyTrpPheThrGlyIleAlaCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGATGGGAAGACATGTGCCAGGCTGAATATCTCTACAAAGAATGTG 319
Db 21 AsnTyrAlaAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAACATGTGATCACTTTCAATGGCTGGCCACACAGCTCCAGTTATCATACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
QY 380 CTTTGGATGAGGAACGGAGTAGGCTGTATGTGGAGCAAGAGTACATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisLePheSerPhe 80
QY 440 GACCTGTTATATCAAGATTTTCAAAGATTGTGTGGCCAGTATCTTACACCAAGA 499
Db 81 AsnLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATGAATCAAGTGGCTGGAAAGACATCCTGAAAGATGTCTAAATTTTCATCAAGTA 559
Db 101 AspGluCysLysTrpPheLysAspIleLeuLysGluCysAlaAsnPhelIleVal 120
QY 560 CTTAAGCATATAATCAGACTCACCTGTACGCTGTGGAACGGGGCTTTTCATCCAAAT 619
Db 121 LeuGluAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATGAATTTGACATCATCTCTGAGGACAAATATTTTAACTGGAGACTCA 679
Db 141 CysThrTyrIleGluValGlyHisHisProGluAspAsnIlePheLysLeuGluAspSer 160
QY 680 CATTTTGAAACGGCGTGGGAAGAGTCCATATGACCTTAAGCTCTCACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGGAATATTACTCTGGAATCGAGCTGATTTTATGGGCGAGACTTTGCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCCGAACCTTTGGGCACACCAACCAATCAGGACAGACAGCATGATTCAGGTGG 859

Db 201 IlePheArgThrLeuGlyAspHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCAATGATCCAAAGTTCAATTAGTGGCCACCTCATCTCAGAGAGTGACAAATCCTCAAGAT 919
Db 221 LeuAsnAspProArgPheIleSerAlaHisLeuIleProGluSerAspAsnProGluAsp 240
QY 920 GACAAAGATATCTTTTCTCCGTGAAATCAATAGATGAGAACACTCTGGAAAAGCT 979
Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleGlyGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTGAGTAAGTGCAGATATGCAGATCACTTTGGAGGCGCAGAGTCTGGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGCAACATTCCTCAAGCTCGTCTGATTGTCTCAGTGCACAGGTCCTCAATGGC 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGTAGTAACTGCAGATGTATTTCTTAATGAATCTTAAAGATCCTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnSerLysAspProLys 320
QY 1160 AATCCAGTTGTATGAGTGTCTTACGACTTCCAGTACATTTTCAAGGGATCAGCCGTG 1219
Db 321 AsnProIleValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGATGATGAGAGGGTGTCTTCTGTGCTCATATGCCACACAGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCAACTATCAATGGTGGCTTTTCAAGGAAGAGTCCCTATCCACGCCAGGAAT 1339
Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCCGCAACAACTTTGGTGGTTTGTACTCTCAAGAGGACCTTCTCATGATGTTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGGCAAGACTCATCCAGCCATGTACAACTCCAGTGTTCCTATCAACAACTGCCCA 1459
Db 401 ThrPheGlyArgSerHisProAlaMetTyrAsnProValPheProIleAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACGGATGTAATTTATCAATTTACAAATTTGCTGAGACCGAGTGGAT 1519
Db 421 IleMetIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAAGTGCAGATGATGATTTATGTTATCGAACAGATGTGGACCGTCTTTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGTGTATGATTTAGAAAGAGTTCCTCGAAGAAATG 1639
Db 461 ValValSerValProLysGluThrTrpHisAspLeuGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTCGGGAACCGACTGTCTATTTCAGCAATGGAGCTTTCCTACAGCAGCAACA 1699
Db 481 ThrValPheArgGluProThrThrIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
QY 1700 CTATATATTGGTTCAACGGCTGGGTTGCCAGCTCCCTTTACACCGGTGTGATTTTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGCTGTGCTGAGTGTGTGCTCGCCGAGACCTTACTGTGCTGGTGGATGGTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCCACTGCAAGAGAGCGCAAGACGACAGATATAGAAT 1879
Db 541 SerCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTCTGTTTCAAGCTTACAC---CATGATATATCATCATGGCCACAGC 1936
Db 561 GlyAspProLeuThrHisCysSerAspLeuGluAspHisAspAsnHisHisGlyProSer 580

Db 381 CysProSerLysThrPheGlyCysPheAspSerThrLysAspLeuProAspGluValIle 400
QY 1400 ACCTTTGCAAGAGTATCCAGCCATGATCAATCCAGTGTTCCTATGAACAATCCGCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyraAsnProValPheProIleAsnSerArgPro 420
QY 1460 ATAGTGATCAAAACCGATGAATATATCAATTTACAAATTTGCTGAGACCGAGTGGAT 1519
Db 421 IleMetIleLysThrAspValAspTyrglnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAGATGACAGTATGATGATTTATTCGGAACAGATGTTGGACCGTCTCTTAAA 1579
Db 441 AlaGluAspGlyGlnTyraAspValMetPheIleGlyThrAspIleGlyThrValIleuLys 460
QY 1580 GPAGTTTCAATTCCTAAGAGACTTGGTATGATTTAGAGAGTTCCTGCTGGAAGAAATG 1639
Db 461 ValValSerIleProLysGlnThrTrpHisGluLeuGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTCGGGAACCGACTCTATTTCCAGCAATGAGAGCTTCCACTAAGCAGCAACA 1699
Db 481 ThrValPheArgGluProThrValIleSerAlaMetLysIleSerThrLysGlnGlnGln 500
QY 1700 CTATATATTGGTTCACCGCTGGGCTGCCAGCTCCCTTTACACCGGTGTGATTTTAC 1759
Db 501 LeuIleGlySerAlaThrGlyValSerGlnLeuProLeuHisArgCysAspValTy 520
QY 1760 GCGAAGCGGTGTGCTGAGTGTTCCTCCGAGACCTTACTGTGCTGGATGGTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyraAspIlePheAspGlySer 540
QY 1820 GCATGTTCTCGTATTTCCCACTCCAAAGAGACGCAACAGACGCAAGATATAGAAT 1879
Db 541 SerCysSerArgTyraPheProThrAlaLysArgArgThrArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTCTGTTTCACACTTA---CACCATGATTAATCACCATGGCCACAGC 1936
Db 561 GlyAspProLeuThrHisCysSerAspLeuGlnHisHisAspAsnProSerGlyGlnThr 580
QY 1937 CCTGAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGACATTTTTCGAATCGAGTCCG 1996
Db 581 LeuGluGluLysIleIleTyrglyValGluAsnSerSerThrPheLeuGluCysSerPro 600
QY 1997 AAGTCGAGAGAGCGCTGCTATTTGGCAATTCGAGAGCGGAATGAAGAGCGGAAAGAA 2056
Db 601 LysSerGlnArgAlaIleValTyrrpGlnPheGlnLysGlnAsnAspAspHisLysVal 620
QY 2057 GAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGCGCTTCTCTACGTAGTCTA 2116
Db 621 GluIleLysValAspAspArgMetIleArgThrGluGlnGlyLeuLeuLeuArgSerLeu 640
QY 2117 CAACAGAAGGATTCAGGCAATTAACCTCTGCCATGGGTGGAAACATGGGTTCATACAAACT 2176
Db 641 GlnArgAspSerGlyIleTyraPheCysHisAlaValGluHisGlyPheIleGlnThr 660
QY 2177 CTTCTTAAGGTAAACCTCGAAGTATTGACACAGAGCAATTTGGAAGAACTTCTTCTATAAA 2236
Db 661 LeuLeuLysValThrLeuGluValIleAspThrAspHisLeuGluGluLeuHisLys 680
QY 2237 GATGATGATGAGATGGTCTTAAGCAACAAAGAAATGCTCAATAGCATGACACCTAGCCAG 2296
Db 681 GluGluAspAlaAspLysThrLysAspAlaThrAsnSerMetThrProSerGln 700
QY 2297 AAGTCTCGTACAGAGACTTCATGAGCTCATCAACACCCCACTCTCAACACAGATGAT 2356
Db 701 LysIleTyrTyraAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAsp 720
QY 2357 GAGTCTCTGAACAGTTTGGAAAGGACCGGAAACAAACCTCGCGCAAGGCCAGGACAT 2416
Db 721 GluPheCysGluGlnValTyrLysArgAspArgLysGlnArgGlnArgProAlaAsn 740
QY 2417 ACCCCAGGACAGTAAACAATGGAGACCTTACAGAAATATAGAAAGGTAGAAACAGG 2476
Db 741 AlaGlnValAsnThrAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArg 760

QY 2477 AGGACCCACGAATTTGAGAGGCGCACCCAGGAGTGTG 2512
Db 761 ArgThrHisGluPheGluArgAlaProArgSerVal 772

RESULT 4

158169

semaphorin III - mouse (fragment)

C:Species: Mus musculus (house mouse)

C>Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 24-Sep-1999

C:Accession: 158169

R:Messersmith, E.K.; Leonardo, E.D.; Shatz, C.J.; Tessier-Lavigne, M.; Goodman, C.S.; Ko

Neuron 14, 949-959, 1995

A:Title: Semaphorin III can function as a selective chemorepellent to pattern sensory

A:Reference number: 158169; MUID:95267432; PMID:7748562

A:Accession: 158169

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-666 <RES>

A:Cross-references: GB:L40484; NID:g703189; PIDN:AAA73934.1; PID:g703190

C:Genetics:

A:Gene: SemaIII

C:Superfamily: semaphorin

Alignment Scores:

Pred. No.:	1.01e-255	Length:	666
Score:	3478.50	Matches:	633
Percent Similarity:	98.50%	Conservative:	23
Best Local Similarity:	95.05%	Mismatches:	9
Query Match:	71.05%	Indels:	1
DB:	2	Gaps:	1

US-09-774-490-1 (1-2709) x 158169 (1-666)

QY 518 GGAAGAAGACATCTCGAAGAAGATGTCTAATTTTCATCAAGGTACTTAAGGCATATATCAG 577
Db 1 GlyLysAspIleLeuLysGluCysAlaAsnPheIleLysValLeuGluAlaTyrAsnGln 20
QY 578 ACTCAGTGTGACGCTGTGGAACGGGGCTTTTCATCAATTCGACCTACATGAAAT 637
Db 21 ThrHisLeuTyraLysGlyThrGlyAlaPheHisProIleCysThrTyrlleGluVal 40
QY 638 GGACATCATCTGAGGACAATATTTTAAAGCTGGAGAACTCACATTTTGAAGACGGCGT 697
Db 41 GlyHisProGluAspAsnIlePheLysLeuGlnAspSerHisPheGluAsnGlyArg 60
QY 698 GGGAGAGTCCATATGACCCCTAAGCTCTGACACATCCCTTTTAAATAGATGGAATTA 757
Db 61 GlyLysSerProTyraAspProLysLeuLeuThrAlaSerLeuLeuIleAspGlyGluLeu 80
QY 758 TACTCTGAACTGCGAGCTGATTTTATGGGCGAGACTTTTGTATCTTCCGAACCTCTTGGG 817
Db 81 TyrsGlyThrAlaAlaAsnPheMetGlyArgAspPheAlaIlePheArgThrLeuGly 100
QY 818 CACCACCAACCAATCAGGACAGACGATGATTCAGGTGGCTCAATGATCCAAAGTTC 877
Db 101 HisHisProIleArgThrGluGlnHisAspSerArgTrpLeuAsnAspProArgPhe 120
QY 878 ATTAGTCCCACTCTATCTCAGAGAGTGCACATCTCTGAAGATGACAAAGTATATCTTTTC 937
Db 121 IleSerAlaHisLeuIleProGluSerAspAsnProGluAspAspLysValTyraPhePhe 140
QY 938 TTCGCTGAAATGCAATAGATGGAACACCTCTCGGAAAGCTACTCCAGCTAGATAGGT 997
Db 141 PheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAlaThrHisAlaArgIleGly 160
QY 998 CAGATATCAAGATGACTTTTGGAGGCGCACAGAGTCTGGTGAATAATAATGACACATTC 1057
Db 161 GlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuValAsnLysTrpThrPhe 180
QY 1058 CTCAAGCTCGTGTGATTTGCTCAGTCCCGAGGTCCAAATGGCATTTGACATCTATTTGAT 1117
Db 181 LeuLysAlaArgLeuIleCysSerValProGlyProAsnGlyIleAspThrHisPheAsp 200

```
QY 1118 GAAGCTGAGGATGATTCCTAATGACCTTTAAAGATCCTTAAATCCAGTTCGTATATGGA 1177
Db 201 GluLeuGlnAspValPheLeuMetAsnSerLysAspProLysAsnProIleValTyrGly 220
QY 1178 GTGTTTACAGCTTCCAGTAACATTTTCAAGGATCAGCGTGTGTATGTATAGCATGAGT 1237
Db 221 ValPheThrThrSerSerAsnIlePheLysGlySerAlaValCysMetTyrSerMetSer 240
QY 1238 GATGTGAGAGGAGTTCCTGGTCCATATGCCACAGGATGGACCCCAACTATCAATGG 1297
Db 241 AspValArgValLeuLeuGlyProTyrAlaHisArgAspGlyProAsnTyrGlnTrp 260
QY 1298 GTGCCCTTACAGGAGAGTCCCTATCCAGCGCAGGACACTTGTCCAGCAAAACATTT 1357
Db 261 ValProTyrGlnGlyArgValProTyrProArgProGlyThrCysProSerLysThrPhe 280
QY 1358 GGTGGTTTTGACTCTACAAGACCTTCTCGATGATGTTTATAACCTTTGCAAGAGTCAAT 1417
Db 281 GlyGlyPheAspSerThrLysAspLeuProAspAspValIleThrPheAlaArgSerHis 300
QY 1418 CCAGCCATGATCAATCCAGTGTCTTCTATGAACAATCGCCCAATAGTATCAAAACGGAT 1477
Db 301 ProAlaMetTyrAsnProValPheProIleAsnAsnArgProIleMetIleLysThrAsp 320
QY 1478 GTAAATATCAATTTACACAATTTGCTAGACCGAGTGGATGCGAGAGATGCGACGAT 1537
Db 321 ValAsnTyrGlnPheThrGlnIleValValAspArgValAspAlaGluAspGlyGlnTyr 340
QY 1538 GATGTTATGTTTATCGAACAGATGTTGGACCGTCTTAAAGTAGTGTTCATCTCCTAAG 1597
Db 341 AspValMetPheIleGlyThrAspValGlyThrValLeuLysValValSerValProLys 360
QY 1598 GAGACTTGTGATGATTTAGAGAGGTTCTGCTGGAAGAAATGACAGTGTTCGGGAACCG 1657
Db 361 GluThrTrpHisAspLeuGluGlnValLeuLeuGluGluMetThrValPheArgGluPro 380
QY 1658 ACTGCTATTTCCAGTATGAGTCTCCACTAGCAGCACAACACTATATTTGGTCAAGC 1717
Db 381 ThrThrIleSerAlaMetGluLeuSerThrLysGlnGlnGlnLeuTyrIleGlySerThr 400
QY 1718 GCTGGGTTGCCAGCTCCCTTTACACCGTGTGATATTACGGGAACCGTGTGCTGAG 1777
Db 401 AlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyrGlyLysAlaCysAlaGlu 420
QY 1778 TGTGCTCTCCCGAGACCTTACTGTGCTGCTGGATGGTTCGCAATGTTCTGCTATTTT 1837
Db 421 CysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySerSerCysSerArgTyrPhe 440
QY 1838 CCCACTGCAAGAGCGCACAGCAGCAAGATATTAAGATGGAAGCCACTGACTCAC 1897
Db 441 ProThrAlaLysArgArgThrArgArgGlnAspIleArgAsnGlyAspProLeuThrHis 460
QY 1898 TGTTCAGACTTA--CACCATGATATCACCATGGCCACAGCCCTGAAGAGAGATCATC 1954
Db 461 CysSerAspLeuGlnHisHisAspAsnHisHisGlyProSerLeuGluGluArgIleIle 480
QY 1955 TATGTTGATAGAAATAGTACCATTTTGGAAATCGAGTCCGAAAGTCGAGAGCGCTG 2014
Db 481 TyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLysSerGlnArgAlaLeu 500
QY 2015 GTCTATTGCAATTCAGAGCGGAATGAAGAGCGGAAGAGAGATCAGAGTGGATGAT 2074
Db 501 ValTyrTrpGlnPheGlnArgArgAsnGluAspArgLysGluGluIleLysMetGlyAsp 520
QY 2075 CATATCATCAGCAGATCAAGCGCTTCTGCTACGTAGTCTACAACAGAGGATTCAGGC 2134
Db 521 HisIleIleArgThrGluGlnGlyLeuLeuLeuLeuLeuLeuLeuLysLysAspSerGly 540
QY 2135 AATTACCTTCGCATCGGTGGAACATGGGTTCATACAAACTCTTCTTAAAGTAAACCTG 2194
Db 541 AsnTyrLeuCysHisAlaValGluHisGlyPheMetGlnThrLeuLeuLysValThrLeu 560
```

```
QY 2195 GAAGTCATTGACACAGAGCATTTTGAAGAACTTCTTCAATAAGATGATGATGAGATGCG 2254
Db 561 GluValIleAspThrGluHisLeuGluGluLeuHisLysAspAspAspGlyAspGly 580
QY 2255 TCTAAGACCAAGAAATGTCCAATAGCATGACACTAGCCAGAAAGTCTGTGTACAGAGAC 2314
Db 581 SerLysIleLysGluMetSerSerSerMetThrProSerGlnLysValTyrTyrArgAsp 600
QY 2315 TTCATGAGCTCATCAACCCACCCCAATCTCAACAGCATGGATGATTTGTGTCAACAAGTT 2374
Db 601 PheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGluPheCysGluGlnVal 620
QY 2375 TGGAAAAGGGACCGAAACCAACGCTCGGCAAGGCCAGGACATACCCCAAGCAACAGTAAC 2434
Db 621 TrpLysArgAspArgLysGlnArgGlnArgProGlyHisSerGlnGlySerSerAsn 640
QY 2435 AAATGGAAGCACTTACAGAAATAAGAAAGGTAGAAACAGGAGGACCCACGAATTTGAG 2494
Db 641 LysTrpLysHisMetGlnGluSerLysLysGlyArgAsnArgThrHisGluPheGlu 660
QY 2495 AGGCGACCCAGGAGTGTTC 2512
Db 661 ArgAlaProArgSerVal 666

RESULT 5
G01856
Semaphorin V - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 24-Sep-1999
C:Accession: G01856
R: Sekido, Y.
submitted to the EMBL Data Library, June 1995
A:Reference number: G08634
A:Accession: G01856
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-749 <SEK>
A:Cross-references: EMBL:U28369; NID:g974283; PIDN:AAD09138.1; PID:g974284
C:Superfamily: semaphorin

Alignment Scores:
Pred. No.: 1,68e-152 Length: 749
Score: 2118.00 Matches: 404
Percent Similarity: 67.22% Conservative: 123
Best Local Similarity: 51.53% Mismatches: 185
Query Match: 43.26% Indels: 72
DB: 2 Gaps: 12

US-09-774-490-1 (1-2709) x G01856 (1-749)
QY 221 GTCTGCTCTTTCTGGGAGTATTACTTACAGCAAGCAAACTATCAGATCGGAGAAC 280
Db 13 LeuAlaLeuLeuTrpAlaValGlyLeuGlySerAlaAla-----Pro 26
QY 281 AATGTGCCAAGCTGAAATATCTTACAAAGAAATGTTGGAATCCAAACAATGTGTACT 340
Db 27 SerProProArgLeuArgLeuSerPheGlnGluGlnAlaTrpHisGlyLeuGlnThr 46
QY 341 TTCAATGCTTGGCCCAACAGCTCCAGTTTATCATACCTTCTCTTTGGATGAGAGACGAGT 400
Db 47 PheSer---LeuGluArgThrCysCysTyrGlnAlaLeuLeuValAspGluGluArgGly 65
QY 401 AGGCTGATATGTTGAGCAAGGATCACAATATTTTCATTCGACCTGGTTAATATC---AAG 457
Db 66 ArgLeuPheValGlyAlaGluAsnHisValAlaSerLeuAsnLeuAspAsnIleSerLys 85
QY 458 GATTTTCAAAAGATTGTGTGGCCAGTATCTTACACCAGAGAGATGAATCAAGTGGGCT 517
Db 86 ArgAlaLysLysLeuAlaTrpProAlaProValGluTrpArgGluGluCysAsnTrpAla 105
QY 518 CGAAAGACATCTCGAAAGAAATGTCTAATTTCAATCAAGGACTTAAAGCATATATCAG 577
Db 106 GlyLysAspIleGlyThrGluCysMetAsnPheValLysLeuLeuHisAlaTyrAsnArg 125
```

QY 578 ACTCACTGTAGCCCTGCTGGAGCGGGGGCTTTTCATCCAAATTTGCACCTACACATGAAATTT 637
 Db ThrHisLeuLeuAlaCysGlyThrGlyAlaPheHisProThrCysAlaPheValGluVal 145
 QY 638 GCACATCATCTCTGAGGACAATATTTTAAAGCTGGAGAACTCACATTTTGAACACGGCGT 697
 Db GlyHisArgAlaGluGluProValLeuArgLeuAspProGlyArgIleGluAspGlyLys 165
 QY 698 GGGAGAGTCCATATGACCTAAGCTGTCTGACGACATCCCTTTTAAATAGATGAGAAATTA 757
 Db GlyLysSerProTyrAspProArgHisArgAlaAlaSerValLeuValClyGluGluLeu 185
 QY 758 TACTCTGGNACTCGAGCTGATTTTATGGGCGAGACTTTTGCTATCTTCGGAACCTCTGGG 817
 Db TyrSerGlyValAlaAlaAspLeuMetGlyArgAspPheThrIlePheArgSerLeuGly 205
 QY 818 CACCACCCCAATCAGGACAGCAGCATGATTCAGGTGGCTCAATGATCCCAAGTTC 877
 Db GlnArgProSerLeuArgThrGluProHisAspSerArgTrpLeuAsnGluProLysPhe 225
 QY 878 ATTAGTGCCCACTCATCTCAGAGATGACATCTCTGAAGATGACAAAGTATACATTTTTC 937
 Db ValLysValPheTrpIleProGluSerGluAsnProAspAspLysIleTrpPhePhe 245
 QY 938 TTCCTGAAATGCAATAGATGGAGAACACTCTCT---GGAAAGCTACTCACGCTTGAATA 994
 Db PheArgGluThrAlaValGluAlaAlaProAlaLeuGlyArgLeuSerValSerArgVal 265
 QY 995 GGTCAATATGCAAGAATGACTTTGGGGGCGACAGACTCTGGTGAATAAATGGACACA 1054
 Db GlyGlnIleCysArgAsnAspValGlyGlnArgSerLeuValAsnLeuTrpThrThr 285
 QY 1055 TTCTCAAAAGCTCGTCTGATTTGCTCAGTGCCAGGTCCCAATGGCATTTGACACTCATTTT 1114
 Db PheLeuLysAlaArgLeuValCysSerValProGlyValGluGly---AspThrHisPhe 304
 QY 1115 GATGAACCTGCGAGTATATCTTAATGAATCTTAAAGATCCCTAAATAATCCAGTTGTATAT 1174
 Db AspGlnLeuGlnAspValPheLeuLeuSerSerArgAspHisArgThrProLeuLeuTyr 324
 QY 1175 GGAGTGTTTACGACTTCCAGTAACATTTTCAAGGATCAGCGGTGTGTATGTATAGCATG 1234
 Db AlaValPheSerSerSerSerIlePheGlnGlySerAlaValCysValTyrSerMet 344
 QY 1235 AGTGATGTGAGAAGGTGTCTCTGGTCCATATGCCACAGGTGGACCCCAACTATCAA 1294
 Db AsnAspValArgAlaPheLeuGlyProPheAlaHisLysGluGlyProMetHisGln 364
 QY 1295 TGGTGCCCTTATCAAGGAAGAGTCCCTTATCCAGCGCCAGGAACCTTGTCCAGCAAAACA 1354
 Db TrpValSerTyrGlnGlyArgValProTyrProArgProGlyMetCysProSerLysThr 384
 QY 1355 TTTGTGTGTTTGTACTCTACAAAGACCTTCTGTATGATGTATTAACCTTTGCCAGAGT 1414
 Db PheGlyThrPheSerSerThrLysAspPheProAspValIleGlnPheAlaArgAsn 404
 QY 1415 CATCAGCCATGTACAAATCCAGTGTCTCTATGAACATCCGCCAATAGTGTATCAAAACG 1474
 Db HisProLeuMetTyrAsnSerValLeuProThrGlyGlyArgProLeuPheLeuGlnVal 424
 QY 1475 GATGTAATTAATTAATTTACAAATTTCTGTAGACCGAGTGGATGCGAGAGATGGACAG 1534
 Db GlyAlaAsnTyrThrPheThrGlnIleAlaAlaAspArgValAlaAlaAspGlyHis 444
 QY 1535 TATGATGTTATGTTTATCGGACAGACTTCTGGGACCGTCTTAAAGTATGTTTCAATTCCT 1594
 Db TyrAspValLeuPheIleGlyThrAspValGlyThrValLeuLysValIleSerValPro 464
 QY 1595 AAGGAGACTTGTGATGATTTAGACAGGTTCTCTGGGAAGAAATGACAGTTTTCGGGAA 1654
 Db LysGlySerArgProSerAlaGluGlyLeuLeuLeuGluGluLeuHisValPheGluAsp 484

QY 1655 CCGAGCTGCTATTTCAGCAATGAGAGCTTTCCACTAAGCAGCAACCACTATATATTGTTCA 1714
 Db SerAlaAlaValThrSerMetGlnIleSerSerLysArgHisGlnLeuTyrValAlaSer 504
 QY 1715 ACGCTGGGGTGGCCAGCTCCCTTTACACGGGTGTATATTACCGGAAACGGTGTGCT 1774
 Db ArgSerAlaValAlaGlnIleAlaLeuHisArgCysAlaAlaHisGlyArgValCysThr 524
 QY 1775 GAGTGTTCCTCGCCCGCAGACCTTACTGTCTGCTGGGATGGTTCGTCTCTCTCGCTAT 1834
 Db GluCysCysLeuAlaAlaArgAspProTyrCysAlaTrpAspGlyValAlaCysThrArgPhe 544
 QY 1835 TTTCCACTGCAAGAGACGACACAGCAAGATATAAGAAATGGAGACCCACTGACT 1894
 Db GlnProSerAlaLysArgPheArgArgGlnAspValArgAsnGlyAspProSerThr 564
 QY 1895 CACTGTTCAGACTTACACCATGATAATCACCATGCCACGCCCTGAA----- 1942
 Db LeuCysSer-----GlyAspSerSerArgProAlaLeuLeu 576
 QY 1943 GAGAGAATCATCTATGTGTAGAGAATAGTAGCACATTTTGGATGCGAGTCCGAGATCG 2002
 Db GluHisLysValPheGlyValGluGlySerSerAlaPheLeuGluCysGluProArgSer 596
 QY 2003 CAGAGAGCGTGTCTATTGGCAATTCAGAGGCGCAATGAAGAGCGCAAAAGAGATC 2062
 Db LeuGlnAlaArgValGluTrpThrPheGlnArgAlaGlyValThrAlaHisThrGlnVal 616
 QY 2063 AGAGTGGATGATCATATCATCAGCAGCAGATCAAGCCCTTCTGCTACGTAGTCTCAACAG 2122
 Db LeuAlaGluGluArgThrGluArgThrAlaArgGlyLeuLeuLeuArgLeuArgArg 636
 QY 2123 AAGGATTCAGGCAATTTACCTCTGCGTCGCGTGAACATGGTTCATACAACTCTTCTT 2182
 Db ArgAspSerGlyValTyrLeuCysAlaAlaValGluGlnGlyPheThrGlnProLeuArg 656
 QY 2183 AAGTAAACCTTGGAAAGTCATTGACACAGAGCATTTGGAAGAACTTCTTATAAAGATGAT 2242
 Db ArgLeuSerLeuHisValLeuSerAlaThrGlnAlaGluArgLeu----- 671
 QY 2243 GATGAGATGGCTCTAAGACCAAGAAATGTCCAAATAGCATGACACCTAGCCAGAGATC 2302
 Db -----AlaArgAlaGluGluAlaAlaProAlaAlaProGlyProLysLeu 687
 QY 2303 TGGTACAGAGACTTCATGTCAGTCTCATCAACCCCACTCTCAACACGATGATGAGTTC 2362
 Db TrpTyrArgAspPheLeuGlnLeuVal----- 696
 QY 2363 TGTGAACAAGTTTGGAAAAGGGACCGAAACAAACGTCGCGCAAGGCCAGGACATACCCCA 2422
 Db -697-----GluProGly---GlyGly 701
 QY 2423 GGGAAACAGTAACAATGGAAG-----CACTTA 2449
 Db ----- 721
 QY 2450 CAAGAAAATGAAGAGGTAGAAACAGGAGGACCCAC-----GAATTTGAGAGG 2497
 Db ----- 741
 QY 2498 GCACCCAGGAGT 2509
 Db GlyProArgSer 745

RESULT 6

148744

Semaphorin A - mouse

C:Species: Mus musculus (house mouse)

C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999

C/Accession: 148744

R:Busch, A.W.; Adams, R.H.; Betz, H.

A>Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates

A:Reference number: I48744; MUID:95267431; PMID:7748561
A:Accession: I48744
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-748 <RES>
A:Cross-references: EMBL:X85990; NID:g854323; PIDN:CAAS9982.1; PID:g854324
C:Genetics:
A:Gene: sema
C:Superfamily: semaphorin

Alignment Scores:
Pred. No.: 4,62e-148 Length: 748
Score: 2059.50 Matches: 400
Percent Similarity: 66.93% Conservative: 112
Best Local Similarity: 52.29% Mismatches: 184
Query Match: 42.06% Indels: 69
DB: 2 Gaps: 15

US-09-774-490-1 (1-2709) x I48744 (1-748)

```
QY 281 AATGTCGAAGCTGAATATTCCTACAAAGAAATGTTGGAAATCCAAACATGTGATCACT 340
DB 27 AsnLeuProArgLeuArgLeuSerPheGlnGluLeuGlnAlaArgHisGlyValArgThr 46
QY 341 TTCAATGGCTTGGCCAAAGCTCCAGTTATCATACCTCTCTTTGGATGAGGAACGGAGT 400
DB 47 PheArg---LeuGluArgThrCysCysTyrGluAlaLeuLeuValAspGluArgGly 65
QY 401 AGGCTGTATGTTGGCAAGGATCACATATTTTCATTCGACCTGGTTAATATC---AAG 457
DB 66 ArgLeuPheValGlyAlaGluAsnHisValAlaSerLeuSerLeuAspAsnHisSerLys 85
QY 458 GATTTTCAAAAGATGTGTGGCCAGTATCTTACACAGAGAGATGAATGCAAGTGGGCT 517
DB 86 ArgAlaLysLysLeuAlaTrpProAlaProValGluTrpArgGluGluCysAsnTrpAla 105
QY 518 GGAAGAAGACATCTGAAGAAGTGTGCTAATTTTCATCAAGGTACTTAAGCATATATCAG 577
DB 106 GlyLysAspIleGlyThrGluCysMetAsnPheValArgLeuLeuHisAlaTyrAsnHis 125
QY 578 ACTCACTTGTACGCTGTGGACGGGGCTTTTCATCCAAATTCG-----622
DB 126 ThrHisLeuLeuAlaCysArgThrGlyAlaPheHisProThrCysAlaLeuTrpArgTrp 145
QY 623 -----ACCTACATTGAAATGGACATCATCTCGAGGACAAATATTTTAAG 667
DB 146 AlaThrAlaGlyGlyThrHisAlaSerThrGly-----ProGlu-----158
QY 668 CTGGAGAATCAATTTTCAAAACGGCGCTGGGAAGAGTCCATATGACCCCTAAGCTGCTG 727
DB 159 -----LysLeuGluAspGlyLysGlyLysThrProTyrAspProArgHisArg 174
QY 728 ACAGCATCCCTTTTAATAGATGAGAAATATATCTCTGGAACCTGCAGCTGATTTATGGGG 787
DB 175 ProProSerValLeuValGlyGluGluLeuTyrSerGlyValThrAlaAspLeuMetGly 194
QY 788 CGAGCTTTGCTATCTTCGAACTCTTGGGCACCCACCACCAATCAGGACAGAGCAT 847
DB 195 ArgAspPheThrIlePheArgSerLeuGlyGlnAsnProSerLeuArgThrGluProHis 214
QY 848 GATTCAGGTGGCTCAATCATCAAAAGTTCATTAGTGGCCCACTCATCTCAGAGATGAC 907
DB 215 AspSerArgTrpLeuAsnGluProLysPheValLysValPheTrpIleProGluSerGlu 234
QY 908 AATCTCGAAGATGACAAAGTATATCTTTCTCCGTGAAAATGCAATAGATGGGAACAC 967
DB 235 AsnProAspAspAspLysIleTyrPhePheArgGluSerAlaValGluAlaPro 254
QY 968 TCT---GGAAAAGCTACTCAGCTAGAAATAGGTACAGATATGCAAGATCACTTTGAGGG 1024
DB 255 AlaMetGlyArgMetSerValSerArgValGlyGlnIleCysArgAsnAspLeuGlyGly 274
QY 1025 CACAGAAGTCTGGTGAATAATGGACAACATTCCTCAAAGCTCGTGTGATTTGCTCAGTG 1084
```

```
DB 275 GlnArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeuValCysSerVal 294
QY 1085 CCAGTCCAAATGGCATTCAGACTCACTTTCATGAACTGCAGGATGTATTCCTATGAAC 1144
DB 295 ProGlyValGluGly---AspThrHisPheAspGlnLeuGlnAspValPheLeuLeuSer 313
QY 1145 TTAAAGATCCTAAATAATCCAGTTGTATATCGAGTGTTCAGACTTCAGTAACATTTTC 1204
DB 314 SerArgAspArgGlnThrProLeuLeuTyrAlaValPheSerThrSerSerGlyValPhe 333
QY 1205 AAGGATCAGCCGTGTGTATATAGCATGATGTGAGAAGGGTGTTCCTTGGTCCA 1264
DB 334 GlnGlySerAlaValCysValTyrSerMetAsnAspValArgAlaPheLeuGlyPro 353
QY 1265 TATGCCACAGGGATGAGCCCACTCAATCAATGGTGGCTTATCAAGGAAGAGTCCCTAT 1324
DB 354 LeuProHisLysGluGlyProThrHisGlnTrpValSerTyrGlnGlyArgValProTyr 373
QY 1325 CCAGGCCAGAACTTGTCCAGCAAAACATTTGGTGGTTCCTCTACAAAGGACCTT 1384
DB 374 ProArgProGlyMetCysProSerLysThrPheGlyThrPheSerSerThrLysAspPhe 393
QY 1385 CCTGATGATGTTATAACCTTTTCAAGAGTTCATCCAGCCATGTACAATCCAGTGTTCCT 1444
DB 394 ProAspAspValIleGlnPheGlyArgAsnHisProLeuMetTyrAsnProValLeuPro 413
QY 1445 ATGAACATCCCAATAGTATCAAAACGGATGTAAATATATCAATTTTACACAAATGTC 1504
DB 414 MetGlyGlyArgProLeuPheLeuGlnValGlyAlaGlyTyrThrPheThrGlnIleAla 433
QY 1505 GTAGACCGAGTGGATGACAGAGATGGACAGATGATGATGTTATGTTTATCGGAACAGATGT 1564
DB 434 AlaAspArgValAlaAlaAspLysHisTyrAspValLeuPheIleGlyThrAspVal 453
QY 1565 GGGACCGTCTTAAGTAGTTCATTCCTAAGGAGACTTCGTATGATTTAGAAAGAGTT 1624
DB 454 GlyThrValLeuLysValIleSerValProLysGlyArgArgProAsnSerGluGlyLeu 473
QY 1625 CTGCTGGAAGAAATGACAGATTTTTCGGGAACCGACTGCTATTTTACGCAATGGAGCTTCC 1684
DB 474 LeuLeuGluGluLeuGlnValPheGluAspSerAlaAlaIleThrSerMetGlnIleSer 493
QY 1685 ACTAAGCAGCAACATATATATGTTCAACGGCTGGGTTCGCCAGCTCCCTTTTACAC 1744
DB 494 SerLysArgGlnGlnLeuTyrValAlaSerArgAlaAlaValAlaGlnIleAlaLeuHis 513
QY 1745 CGGTGTGATATTTACGGAAAGCGTGTGCTGAGTGTTCCTCGCTCGCCGAGACCTTACTGT 1804
DB 514 ArgCysThrAlaLeuGlyArgAlaCysAlaGluCysCysLeuAlaArgAspProTyrCys 533
QY 1805 GCTTGGGATGTTCTGCATGTTCTCGCTATTTTCCACTGCAAGAGAGCCACAGCA 1864
DB 534 AlaTrpAspGlySerAlaCysThrArgPheGlnProThrAlaLysArgArgPheArg 553
QY 1865 CAAGATATAAGAAATGGAGACCCCTGACTCCTCTGCTTCAGACTTACACCATGATATCAC 1924
DB 554 GlnAspIleArgAsnGlyAspProSerThrLeuCysSerGly-----AspSerSer 570
QY 1925 CATGCCACACCCCTTGAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGCATTGTTG 1984
DB 571 HisSerValLeuLeuGluLysLysValLeu---GlyValGluSerGlySerAlaPheLeu 589
QY 1985 GAATGCACTCGAAGTCCAGAGAGCGCTGTGCTATTTGGCAATTCAGAGGGGAAATGNA 2044
DB 590 GluCysGluProArgSerLeuGlnAlaHisValGlnTrpThrPheGlnGlyValGlyGlu 609
QY 2045 GAGCGAAAGAGAGATCAGATGATCATATCATCATCAGCAGAGATCAAGCCCTTCTG 2104
DB 610 AlaAlaHisThrGlnValLeuAlaGluGluArgValGluArgThrAlaArgGlyLeuLeu 629
QY 2105 CTACGTAGTCTACAAACAGAGAGATTTACGCAATTTACCTCTGCCATCGCGTGGAAATGGG 2164
```


Db 445 ArgTyrGluValLeuPheLeuGlyThrAspArgGlyThrValGlnLysValleValLeu 464
QY 1592 CTAAGAGAGACTTGGTATGATTAGAGAGGTTCTGCTGGAAGAATGACAGATTTTCGG 1651
Db 465 ProLysAspAsp---GlnGluMetGluGluLeuMetLeuGluValGluValPheLys 483
QY 1652 GAACGACTGCTATTTCACATGAGCTTCCACTACGACACCACTATATATGGT 1711
Db 484 AspProAlaProValLysThrMetThrIleSerSerLysArgGlnGlnLeuValAla 503
QY 1712 TCAACGGCTGGGTTCCCGAGCTCCCTTTACACCGGTGTATATTTACGGGAAGCGTGT 1771
Db 504 SerAlaValGlyValThrHisLeuSerLeuHisArgCysGlnAlaTyrGlyAlaCys 523
QY 1772 GCTGAGTGTGCTCCCGAGACCTTACTGTGCTGGATGGTGTCTGCTGCTGCTG 1831
Db 524 AlaAspCysCysLeuAlaArgAspProTyrCysAlaTyrAspGlyGlnAlaCysSerArg 543
QY 1832 TATTTTCCCACTGCAAGAGAGCGCACAGACGACAGATATAGAAATGAGACCACTG 1891
Db 544 TyrThrAlaSerSerLysArgArgSerArgGlnAspValArgHisGlyAsnProIle 563
QY 1892 ACTCACTGTTCAGACTTACACCATGATATACCATGGCCACAGCCCTGAGAGAGATC 1951
Db 564 ArgGlnCysArgGlyPhe-----AsnSerAsnAlaAsnLysAsnAlaValGluSerVal 581
QY 1952 ATCTATGTTGAGAAATAGTAGACATATTTTGGAAATGAGTCCGAGTCCGAGAGAGG 2011
Db 582 GlnTyrGlyValAlaGlySerAlaAlaPheLeuGluCysGlnProArgSerProGlnAla 601
QY 2012 CTGCTATTGCAATTCAGAGGCGGAATGAAGCGGAAAGACAGACATGAGTGGAT 2071
Db 602 ThrValLysThrLeuPheGlnArgAspProGlyAspArgArgGlnLysLeuAlaGln 621
QY 2072 GATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAGTCTACAAAGAGATTC 2131
Db 622 AspArgPheLeuArgThrGluGlnGlyLeuLeuLeuArgAlaLeuGlnLeuSerAspArg 641
QY 2132 GGCATTTACTCTGCGATCGGTGGAAATGAGTTCATCAAACTCTTCTTAAAGTAAAC 2191
Db 642 GlyLeuTyrSerCysThrAlaThrGluAsnAsnPheLysHisValValThrArgValGln 661
QY 2192 CTGGAATCATTCACACAGACATTTGGAAGAACTCTTCAATAAGATCATGATGGAT 2251
Db 662 LeuHisValLeuGlyArgAspAlaValHisAlaLeuPhe----- 675
QY 2252 GGTCTTAAGACCAAGAAATGCTCAATAGCATGACACCTAGCCAGAGGCTCTGG----- 2305
Db 676 -----ProLeuSerMetSerAlaProProProGlyAlaGlyProPro 691
QY 2306 -----TACAGAGACTTCATGACGCTCATCAACCCCAATCTCAACAGATGGAT 2356
Db 692 ThrProProTyrGlnGluLeuAlaGlnLeuLeuAlaGlnProGluValGlyLeuLeuHis 711
QY 2357 GAGTTCGTGCAACACTTGGAAAGGAGGACCGGAAACACGTCGCGAAAGGCGAGACAT 2416
Db 712 GlnTyrGlnGlyTyrTrpArgHisValProProSerProArgGluAlaProGly---- 730
QY 2417 ACCCCAGGGAACAGTAACAAATGGAAGCACTTACAGAAAAATAAGAAAGGTAGAACAGG 2476
Db 731 -----AlaProArgSerProGluProGlnAspGlnLysLysProArgAsnArg 746
QY 2477 AGGACCCAC 2485
Db 747 ArgHisHis 749

RESULT 8

148748

semaphorin E - mouse

C;Species: Mus musculus (house mouse)

C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999

C;Accession: I48748

R;Fuschel, A.W.; Adams, R.H.; Betz, H.

Neuron 14, 941-948, 1995
A;Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates
A;Reference number: I48744; MUID:95267431; PMID:7748561
A;Accession: I48748
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-751 <RES>
A;Cross-references: EMBL:X85994; NID:9854331; PIDN:CAA59986.1; PID:9854332
C;Genetics:
A;Gene: semE
C;Superfamily: semaphorin

Alignment Scores: 3.31e-132 Length: 751
Pred. No.: 1850.50 Matches: 343
Score: 65.54% Conservative: 161
Percent Similarity: 65.54% Mismatches: 224
Best Local Similarity: 44.60% Indels: 41
Query Match: 37.80% Gaps: 10
DB: 2

US-09-774-490-1 (1-2709) x I48748 (1-751)

QY 215 AGGATTGCTCTCTTTCTGGGAGTATTACTTACAGCAAGCAAACTATCAGATGGG 274
Db 4 ArgAlaIleCysValLeuValGlyValPheIleCysSerIleCys---ValArgGlySer 22
QY 275 AAGAACATGTCGCAAGCTGAAATATCTTACAAAGAAATGTTGGATCCAACTATGTG 334
Db 23 SerGlnProGlnAlaArgValTyrLeuThrPheAspGluLeuArgGluThrLysThrSer 42
QY 335 ATCACTTTCAATGGCTGGCCCAACAGCTCCAGTATCATCTCTCTTTGGATGAGGAA 394
Db 43 GluTyrPheSerLeuSerHisGlnLeuAspTyrArgIleLeuLeuMetAspGluAsp 62
QY 395 CGGAGTAGGCTGTATGTTGGAGCAAGGATCACATATTTTCATTCGACCTGGTTAATATC 454
Db 63 GlnAspArgIleTyrValGlySerLysAspHisIleLeuSerLeuAsnIleAsnIle 82
QY 455 ---AAGGATTTTCAAAGATTGTGGCCATATCTTACACAGAGAGATGAATGCAAG 511
Db 83 SerGlnGluProLeuSerValPheTrpProAlaSerThrIleLysValGluGluCysLys 102
QY 512 TGGGCTGAAAGACATCTCTGAAAGATGCTTAATTTTCATCAAGGTACTTAAAGCATAT 571
Db 103 MetAlaGlyLysAspProThrHisGlyCysGlyAsnPheValArgValIleGlnThrPhe 122
QY 572 AATCAGACTCTTGTAGCCTGTGGAAACGGGGCTTTTCATCCAAATTTGCACCTACATT 631
Db 123 AsnArgThrHisLeuTyrValCysGlySerGlyAlaPheSerProValCysThrTyrLeu 142
QY 632 GAAATTGACATCATCTCTGAGACATATATTTTAAAGCTGGAGAACTCACATTTGAAAAC 691
Db 143 AsnArgGlyArgSerGluAspGlnValPheMetIleAsp---SerLysCysGluSer 161
QY 692 GCGCTGGGAAGAGTCCATATGACCTAAGCTGTGACAGCATCCCTTTTAAATAGATGA 751
Db 162 GlyLysGlyArgCysSerPheAsnProAsnValAsnThrValSerValMetIleAsnGlu 181
QY 752 GAAATTATCTCTGGAACCTGCAAGCTGAATTTTATGGGCGAGACTTTTGTCTATCTTCCGAACT 811
Db 182 GluLeuPheSerGlyMetTyrIleAspPheMetGlyThrAspAlaAlaIlePheArgSer 201
QY 812 CTGGGCAACCAACCCCAATCAGGACAGACGATGATTCAGGTGGCTCATCATCA 871
Db 202 LeuThrLysArgMetGlnLeuArgThrAspGlnHisAsnSerLysTrpLeuSerGluPro 221
QY 872 AAGTTCATTAGTCCCACTCTCATCAGAGAGTGAACAATCTTGAAGATGACAAAGTATAC 931
Db 222 MetPheValAspAlaHisValIleProAspGlyThrAspProAsnAspAlaLysValTyr 241
QY 932 TTTTCTCTCCGTGAAATGCAATAGATGAGAGAACACTCTGGAAGAGCTACTCAGCTAGA 991
Db 242 PhePhePheLysGluArgLeuThrAspAsnAsnArgSerThrLysGlnIleHisSerMet 261

QY	2069	GATGATCATATCATCAGCACAGATCAAGCGCTTCTCTAGGTAGTTCTACACAGAAGGAT	2129
Db	618	AsnGlnArgIleIleAlaThrSerGlnGlyLeuLeuIleAArgSerValGlnAspSerAsp	637
QY	2129	TCAGGCAATTACCTCTGCCATCGCGTGGAACATGGGTTCATACAAACTCTCTCTTAAGGTA	2188
Db	638	GlnGlyLeuTyHisCysIleAlaThrGluAnsSerPhelysGlnThrIleAlaLysIle	657
QY	2189	ACCCTGGAAGTCATTCAGCACAGCATTTTGGAAAGAACTTCTTCATAAAGATGATGAGA	2248
Db	658	AsnPhelysValLeuAspSer-----GluMetValAlaValAlaThrAspLysTrpSerProTrpThr	678
QY	2249	GATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAGAGGTCCTGGTAC	2308
Db	665	-----GluMetValAlaValAlaThrAspLysTrpSerProTrpThr	678
QY	2309	-----AGAGACTTCATGCAGCTCATCAAC	2332
Db	679	TrpAlaGlySerValArgAlaLeuProPheHisProLysAspIleLeuGlyAlaPheSer	698
QY	2333	CACCCCNATCTAACACGATGGATGATGTTCTGTGAACAAGTTTGGAAAAGGGACCGAAAA	2392
Db	699	HisSerGluMetGlnLeuIleAsnGlnTyrCysLysAspThr-----ArgGlnGlnGin	716
QY	2393	CAACGTCGGCAAAGGCCAGGACATACCCCAGGGAACAGATAACAATGGAGCACTTACAA	2452
Db	717	GlnLeuGlyGlnGluProGlnLysMetArgGlyAspTry-GlyLysLeuLysAlaLeuIle	736
QY	2453	GAATAAAGAAAGGTAGAAACAGGAGG	2479
Db	737	AsnSerArgLysSerArgAsnArgArg	745
RESULT 9			
S66498	M-sema F protein precursor - mouse		
C:	Species: Mus musculus (house mouse)		
C:	Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 24-Nov-1999		
C:	Accession: S66498		
R:	Inagaki, S.; Furuyama, T.; Iwahashi, Y.		
F:	FEBS Lett. 370, 269-272, 1995		
A:	Title: Identification of a member of mouse semaphorin family.		
A:	Reference number: S66498; PMID:95385809; PMID:7656991		
A:	Accession: S66498		
A:	Status: preliminary		
A:	Molecule type: mRNA		
A:	Residues: 1-834 <NA>		
A:	Cross-references: EMBU:S79463; NID:g1110598; PIDN:AAB35184.1; PID:g1110599		
C:	Superfamily: semaphorin		
F:	1-21/Domain: signal sequence #status predicted <SIG>		
F:	22-834/Product: M-sema F protein #status predicted <MAT>		
Alignment Scores:			
Pred. No.:	4,28e-65	Length:	834
Score:	966.00	.Matches:	236
Percent Similarity:	51.35%	Conservative:	107
Best Local Similarity:	35.33%	Mismatches:	269
Query Match:	19.73%	Indels:	21
DB:	2	Gaps:	21
US-09-774-490-1 (1-2709) x S66498 (1-834)			
QY	206	TGTTAACTAGGATGCTGCTCTTTCTGGGGATTAATTACTTACAGCAAGACAACTAT	265
Db	8	TrpLeu-----LeuAlaAlaGlyLeuTrpGlyLeuGlyIleGlyAlaGluMetTrpTrp	25
QY	266	CAGAAATGGGAAGCAAAATGTGCCAAGCGCTGAAATATTCCTACAAAGAAATGTTGGAATCC	325
Db	26	-----AsnLeuValProArgLysThrValSerSerGlyGluLeuValThrVal	41
QY	326	AACAATGTGATCACTTTCAATGGCTTGGCCCAACAGCTCAGTTATCATCACTTCCTTTTG	385
Db	42	ValArgPheSerGlnThrGlyIle-----GlnAspPheLeuThrLeuThrLeu	58

386 GATGAGAACGGAGTAGGCTGATGTTGGAGCAAGGATCACATATTTTCATTCCAGCTG 445
Db ThrGluHisSerGlyLeuLeuValGlyAlaArgGluAlaLeuPheAlaPheSerVal 78
446 GTTAATATCAAGGATTTTCAAAAGATTGTGTGCCAGTATCTTACACAGAAGATGAA 505
Db GluAlaLeuGluLeuGlnGlyAlaIleSerTrpGluAlaProAlaGluLeuValGlu 98
506 TGCAGTGGCGTGGAAAAGACATCTCTGAAGAAATGTGCTAATTTTCAATCAAGTACTTAAG 565
Db CysThrGlnLysGlySerAsnGlnThrGluCysPheAsnPheIleArgPheLeuGln 118
566 GCATATAATCAGACTCAGCTGACGCTGTGGAACGGGGCTTTTCATCCAATTTGCACC 625
Db ProTyAsnSerSerHisLeuValCysGlyThrTyAlaPheGlnProLysCysThr 138
626 TACATTGAAATTTGGACATCATCTCGAGCAATATTTTAAAGCTGGAGAACTCACATTTT 685
Db TyrIleAsnMet-----LeuThrPheThrLeuAspArgAlaGluPhe 152
686 GAAACGGCGTGGGAAGAGTCCATATGACCTTAAGCTCTGACAGCATCCCTTTTAATA 745
Db GluAspGlyLysGlyLysCysProTyAspProAlaLysGlyHisThrGlyLeuLeuVal 172
746 GATGAGAAATTTACTCTGGAACCTGAGCTGATTTTATGGGCGAGACTTGTCTATCTTC 805
Db AspGlyGluLeuTySerAlaThrLeuAsnAsnPheLeuGlyThrGluProValIleLeu 192
806 CGAATCTTTGGGCACCAACCAATCAGACAGCAGCATGATTTCCAGGTGGCTCAAT 865
Db ArgTyrMetGlyThrHisIleSerIleLysThrGlu--TyrLeuAlaPheTrpLeuAsn 211
866 GATCCTAAAGTTCTATGTCGCCACCTCATCTGAGAGAGTGACATCT-----GAA 916
Db GluProHisPheValGlySerAlaPheValProGluSerValGlySerPheThrGlyAsp 231
917 GATGACAAAGTATACATTTTCTCCGTGAAATGCAATAGATGAGAGCACTCTGGAAA 976
Db AspAspLysIleTyrPhePheSerGluArgAlaValGluTyrAspCysTyrSerGlu 251
977 GCTACTCAGCTAGATAGTATGATATGCAAGATGACTTTGGAGGGCAGCAAGTCTG 1036
Db GlnValValAlaArgValAlaArgValCysLysGlyAspMetGlyGlyAlaArgThrLeu 271
1037 GTGAATAATGACATCTCTCAAGCTGCTGATTTGCTCAGTGCAGGTCCCAAT 1096
Db GlnLysLysTrpThrPheLeuLysAlaArgLeuValCysSerAlaPro-----288
1097 GGCATTGACACTCATTTTGTATGAACTGAGGATGATTTCTATGAACCTTAAAGATCCT 1156
Db AspTrpLysValTyrPheAsnGlnLeuLysAlaValHisThrLeuArgGlyAlaSerTrp 308
1157 AAAAAATCCAGTTGTATGAGTGTGTATGAGTCTCCAGCTTCCAGTAAACATTTTCAAGGATCAGCC 1216
Db HisAsnThrThrPhePheGlyValPheGlnAlaArgTyrGlyAspMetAspLeuSerAla 328
1217 GTGTGTATGATAGCATGATGATGAGAGGGTGTCTGCTCCATATGCCACAGG 1276
Db ValCysGluTyrGlnLeuGluGlnIleGlnGlnValPheGluGlyProTyrLysGluTyr 348
1277 GATGGACCAACTATCAATGGGTGCCTTATCAAGGAAGAGTCCCTATCCAGGCCAGCA 1336
Db SerGluGlnAlaGlnLysTrpAlaArgTyrThrAspProValProSerProArgProGly 368
1337 ACTTGTGCCAGCAAAACATTT-----GGTGGTTTTGACTCTACAAAGGACCTTCTGAT 1390
Db SerCysIleAsnAsnTrpHisArgAspAsnGlyTyThrSerSerLeuGluLeuProAsp 388
1391 GATGTTATACCTTTTCAAGAAGTCTATCCAGCCATGATACATCCAGTGTTCCTATGAC 1450
Db AsnThrLeuAsnPheIleLysLysHisProLeuMetGluAspGlnValLysProArgLeu 408
1451 AATCGGCCAATAGTGAACCAAGGATGATAAATTATCAATTTTACAAATTTGTCGTAGAC 1510

409 GlyArgProLeuLeuValLysLysAsnThrAsn-----PheThrHisValValAlaAsp 426
1511 CGATGGATGAGATGAGATGGA---CAGTATGATGTTATGTTATCGGACAGATTTGGG 1567
Db ArgValProGlyLeuAspGlyAlaThrTyrThrValLeuPheIleGlyThrGlyAspGly 446
1568 ACCGTTCTTAAGTAGTTTCAATTCCTAAGGAGACTTCGTATGATTTAGAGAGGTTCTG 1627
Db TrpLeuLeuLysAlaValSerLeu-----GlyProTrpIleHis-----Met 460
1628 CTGAAGAAATGACAGATTTT---CGGAAACCGACTGCTATTTTACAGCAATGAGGTTTCC 1684
Db ValGluGluLeuGlnValPheAspGlnGluPro-----ValGluSerLeuValLeuSer 478
1685 ACTAAGCAGCAACATATATATTTGTTCAACGGCTGGGTTGCCACTCCCTTTTACAC 1744
Db GlnSerLysLysValLeuPheAlaGlySerArgSerGlnLeuValGlnLeuSerLeuAla 498
1745 CGGTGTGATATTTACGGGAAACGCTGCTGAGTGTTCCTCGCCGAGAGACCTTACTGT 1804
Db AspCysThrLysTyr---ArgPheCysValAspCysValLeuAlaArgAspProTyrCys 517
1805 GCTTGGGATGTTCTGCAATGTTCTCGCTATTTTCCCACTGCAAAAGAGACGCAACAGCA 1864
Db AlaTrpAsnValAsnThr---SerArgCysValAlaThrThrSerGlyArgSerGlySer 536
1865 -----CAAGATATAAGAAATGGAGACCCACTGACTGCTACTGTTACAGACTTACCAT 1915
Db PheLeuValGlnHisValAlaAsnLeuAspThrSerysMetCys-----551
1916 GATATACCATCCAGCCACAGCCCTGAAGAGAGATC-----ATCTATGCTGTAGAG 1966
Db AsnGlnTyrGlyLysLysValArgSerIleProLysAsnIleThrValValSer 570
1967 AATAGTACACATTTTGGAAATGCAAGTCCGAGTCCGAGAGAGCGCTGCTTATTTGGCAA 2026
Db GlyThrAspLeuValLeuProCysHisLeuSerSerAsnLeuAlaHisAlaHisTrpThr 590
2027 TTC-----CAGAGCGGAAATGAAGAGCGAAAGAGAGATCAGAGTGGATCATATC 2080
Db PheGlySerGlnAspLeuProAlaGlnProGlySerPheLeuTyAspThrGlyLeu 610
2081 ATCAGGACAGATCAAGGCTTCTGCTAGTGTCTACAAACAGAGAGATTTCAGGCAATTAC 2140
Db -----GlnAlaLeuValValMetAlaAlaGlnSerArgHisSerGlyProTyr 626
2141 CTCTGCCATCGGTGGAAACATGGG 2164
Db ArgCysTyrSerGluGluGlnGly 634

RESULT 10

148746
semaphorin C - mouse (fragment)
C.Species: Mus musculus (house mouse)
C.Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 03-Nov-2000
C.Accession: I48746
R.Puschel, A.W.; Adams, R.H.; Betz, H.
Neuron 14, 941-948, 1995
A.Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates
A.Reference number: I48744; MUID:95267431; PMID:7748561
A.Accession: I48746
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: mRNA
A.Residues: 1-782 <RES>
A.Cross-references: EMBL:X85992; NID:g854327; PIDN:CAA59984.1; PID:g854328
C.Genetics:
A.Gene: semC
C.Superfamily: semaphorin
Alignment Scores: 7.21e-57 Length: 782
Pred. NO.: 857.50 Matches: 217

Percent Similarity:	49.77%	Conservative:	114
Best Local Similarity:	32.63%	Mismatches:	261
Query Match:	17.51%	Indels:	73
DB:	2	Gaps:	23

US-09-774-490-1 (1-2709) x 148746 (1-782)			
QY	356	AACAGCTCCAGTTATCATCATCTCTTTTGGATGAGAACGGAGTAGGCTGTATGTTGGA	415
DB	12	AsnIleSerAenTyThrAlaLeuLeuLeuSerGlnAspGlyLysThrLeuTyrValGly	31
QY	416	GCAAGGATCATATATTTTCATTGCAGCTGGTTATATCAAG-----GAT	460
DB	32	AlaArgGluAlaLeuPheAlaLeuAsn---SerAenLeuSerPheLeuProGlyGlyGlu	50
QY	461	TTTCAAAAGATTGTGGCCAGTATCTTACACCAGAGAGATGAATGCAAGTGGCGCTGGA	520
DB	51	TyrGlnGluLeuLeuTyrSerAlaAspAlaAspArgLysGlnGlnCysSerPheLysGly	70
QY	521	AAAGACATCTCTGAAAGAATGTGTAATTTCATCAAGGTACTTAAAGGCATATAATCAGACT	580
DB	71	LysAspProLysArgAspCysGlnAsnTyxIleLysIleLeuLeuProLeuAsnSerSer	90
QY	581	CACTTGTACGCTCTGTGAACGGGGCTTTTCATCCAATTTCACCTACATTGGAATTGGA	640
DB	91	HisLeuLeuThrCysGlyThrAlaAlaPheSerProLeuCysAlaTyxIleHisIleAla	110
QY	641	CATCATCTCGAGGACAATATTTTAAAGCTGGAG-----AATCCACATTTT	685
DB	111	Ser-----PheThrLeuAlaGlnAspGluAlaGlyAsnValIleLeu	124
QY	686	GAATAACGGCGGTGGAGAGTCCATATGACCTTAAGCTGTGACAGCATCCCTTTTAATA	745
DB	125	GluAspGlyLysGlyHisCysProPheAspProAsnPheLysSerThrAlaLeuValVal	144
QY	746	GATGGAGAAATATACTCTGNACTGCAGCTGATTTTATGGGGCAGACTTTGCTATCTTC	805
DB	145	AspGlyGluLeuTyrThrGlyThrValSerSerPheGlnGlyAsnAspProAlaIleSer	164
QY	806	CGAACTCTTGGGCCACCAACCCCAATCAGGACAGCAGCATGATTCAGGTGCTCAAT	865
DB	165	ArgSerGlnSer---SerArgProThrLysThrGluSer---SerLeuAsnTrpLeuGln	182
QY	866	GATCCAAGTTTCATTAGTGGCCCACTCATCTCAGAGAGT---GACAATCTT-----GAA	916
DB	183	AspProAlaPheValAlaSerAlaThrSerProGluSerLeuGlySerProIleGlyAsp	202
QY	917	GATGACAAAGTATACTTTTCTCCGTGMAAATGCAATAGATGAGAACACTCTGGAAAA	976
DB	203	AspAspLysIleTyrPhePhePheSerGluThrGlyGlnGluPheGluPhePheGluAsn	222
QY	977	GCTACTCACGCTAGCAATAGGTGCAGATATGCAAGAATGACTTTGGAGGCGCACAGAGCTG	1036
DB	223	ThrIleValSerArgValAlaArgValCysLysGlyAspGluGlyGlyLysArgValLeu	242
QY	1037	GTGAAATAATGGACAACATCTCTCAAGCTCGTCTGATTTGCTCAGTGCAGAGTCCCAAT	1096
DB	243	GlnGlnArgTrpThrSerPheLeuLysAlaGlnLeuLeuCysSerArgPro---AspAsp	261
QY	1097	GGCATTCGACACTCAATTTTGATGAACATCGAGATGATTTCTCTAATG-----A	1150
DB	262	GlyPhePro-----PheAsnValLeuGlnAspValPheThrLeuAsnProAsnProGln	279
QY	1151	GATCCTAAAAATCCAGTTGTATATGAGTGTGTTACGACTTCC-----AGTAAACATTTTC	1204
DB	280	AspTrpArgLysThrLeuSerIleGlyValPheThrSerGlnTrpHisArgGlyThrThr	299
QY	1205	AAGGGATCAGCGGTGTGTATGATAGCATAGTATGAGAAAGGGGTGTTCTTGGTCCCA	1264
DB	300	GluGlySerAlaIleCysValPheThrMetAsnAspValGlnLysAlaPheAspGlyLeu	319
QY	1265	TATGCCACAGGGATGATGACCCCACTCAATGGTGGTGTATCAAGGAAGAGTCCCTAT	1324

RESULT 11
I48745
semaphori

C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Nov-1999
C:Accession: I48745
R:Puschel, A.W.; Adams, R.H.; Betz, H.
Neuron 14, 941-948, 1995
A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates
A:Reference number: I48744; MUID:95267431; PMID:7748561
A:Accession: I48745
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-760 <RES>
A:Cross-references: EMBL:X85991; NID:g854325; PIDN:CAA59983.1; PID:g854326
C:Genetics:
A:Gene: semB
C:Superfamily: semaphorin

Alignment Scores:
Pred. No.: 1,17e-52 Length: 760
Score: 802.00 Matches: 204
Percent Similarity: 49.32% Conservative: 122
Best Local Similarity: 30.86% Mismatches: 251
Query Match: 16.38% Indels: 84
DB: 2 Gaps: 23

US-09-774-490-1 (1-2709) x I48745 (1-760)

QY	257	GMAACTATCAGAAATGGGAAGAACAAATGCGCAAGCTGAAATATCTTACAAAGAAATG	316
DB	30	AlaSerGlyThrGlyGlnGlyProMetProArgValLysTyr-----HisAla	46
QY	317	TTGGAATCCAAATGTGATCACTTCAATGGCTTGGCCCAACAGCTCCAGTTATCATACC	376
DB	47	GlyAspGlyHisArgAlaLeuSerPhePheGlnGlnLysGlyLeuArgAspPheAspThr	66
QY	377	TTCTTTTGGATGAGGAACGAGTAGCTGTATGTTGGAGCAAGATCACATATTTTCA	436
DB	67	LeuLeuLeuSerAspAspGlyAsnThrLeuTyrValGlyAlaArgGluThrValLeuAla	86
QY	437	TTGACCTGGTAAAT-----ATCAAGATTTTCAAAAGATTGTGTGGCCA	481
DB	87	LeuAsnIleGlnAsnProGlyIleProArgLeuLysAsnMet-----IleProTrpPro	104
QY	482	GTATCTTACACCAAGAGATGAATCAAGTGGCTGGGAAAGACATCTCGAAAGAACT	541
DB	105	AlaSerGluArgLysLysThrGluCysAlaPheLysLysSerAsnGluThrGlnCys	124
QY	542	GCTAATTTTCATCAAGTACTTAAGGCATATAATCAGACTCAGCTGTACGCTGTGGAACG	601
DB	125	PheAsnPheIleArgValLeuValSerTyrAsnAlaThrHisLeuTyrAlaCysGlyThr	144
QY	602	GGGGCTTTTCATCAATTTGCACCTACATGAATTTGGACATCATCTCGAGGACAAATTT	661
DB	145	PheAlaPheSerProAlaCysThrPheIleGluLeu-----GlnAspSerLeuLeu	161
QY	662	TTTAACTGGAGAATCACATTTTGAAGCGCGGTGGGAAGTCCATATGACCCCTAAG	721
DB	162	LeuProIleLeuIleAspLysValMetAspGlyLysGlnSerProLeuThrLeuPhe	181
QY	722	CTGCTGACGACATCCCTTTAATAGATGGAGAATTATCTCGAACTGAGCTGATTTT	781
DB	182	ThrSerThrGlnAlaValLeuValAspGlyMetLeuTyrSerGlyThrMetAsnAsnPhe	201
QY	782	ATGGGGCGAGACTTCTGCTATCTCCGAATCTTGGGCACACACCACCAATCAGGACAG	841
DB	202	LeuGlySerGluProIleLeuMetArgThrLeuGlySerHisProValLeuLysThrAsp	221
QY	842	CAGCATGATTCAGGTGGCTCAAT---GATCCAAAGTTTCAATAGTCCACCATCTCTCA	898
DB	222	IlePhe---LeuArgTrpLeuHisAlaAspAlaSerPheValAlaAla-----	236
QY	899	GAGAGTGACAATCTCGAAGATGACAAAGTATATCTTTCTTCGCGTGAATGCAATAGAT	958
DB	237	-----IleProSerThrGlnValValTyrPhePhePheGluThrAlaSerGlu	253

QY	959	GGAGAACACTCTGGAAAGCTACTCACGCTAGAATAGTTCAGATATGCAAGATGACTTT	1018
DB	254	PheAspPhePheGluGluLeuTyrIleSerArgValAlaGlnValCysLysAsnAspVal	273
QY	1019	GGAGGGCACAGAAAGTCTGGTGAATAAATGACAAATCTCTCAAGCTGCTGTGATTTGC	1078
DB	274	GlyGlyGluLysLeuLeuGlnLysLysTrpThrThrPheLeuLysAlaGlnLeuLeuCys	293
QY	1079	TCAGTCCAGGTCCAAATGGCATTGACACTCATTTTGTATGAACACTGCAGGATGATTCTTA	1138
DB	294	AlaGlnProGly-----GlnLeuProPheAsnIleIleArgHisAlaValLeu	309
QY	1139	ATGAACCTTAAAGATCTTAAATCCAGTTGTATATGAGTGTCTTACGATCTCCAGTAAC	1198
DB	310	LeuProAlaAspSerProSerValSerArgIleTyrAlaValPheThrSerGlnTrpGln	329
QY	1199	ATT-----TTCAAGGGATCAGCCGTGTGTATATAGCATGAGTGTGAGAGGGTGT	1252
DB	330	ValGlyGlyThrArgSerSerAlaValCysAlaPheSerLeuThrAspIleGluArgVal	349
QY	1253	TTCTTGTGTCATATGCCACAGGATGGACCCAACTATCAATGGTGGCTTATCAAGA	1312
DB	350	PheLysGlyLysTyrLysGluLeuAsnLysGluThrSerArgTrpThrThrTyrArgGly	369
QY	1313	AGAGTCCCTATCCAGCCGACCAACTTGT-----CCAGCAAAACATTTGGTGT	1363
DB	370	SerGluValSerProArgProGlySerCysSerMetGlyProSerSer-----	385
QY	1364	TTTGACTCTCAAGGACCTTCTCTGATGATTTATAACCTTTGCAAGAGTATCATCAGCC	1423
DB	386	-----AspLysAlaLeuThrPheMetLysAspHisPheLeu	397
QY	1424	ATGTACAATCCAGTGTTCCTATGAAACAATCGCCCAATAGTGTCAAAACGGATGTAAT	1483
DB	398	MetAspGluHisValVal-----GlyThrProLeuLeuValLysSerGlyValGlu	414
QY	1484	TATCAATTTACAAATTTGTCGACCGAGTGGATGCAGAGATGGA-----CAGTATGAT	1540
DB	415	Tyr-----ThrArgLeuAlaValGluSerAlaArgGlyLeuAspGlySerSerHisVal	432
QY	1541	GTATGTTTATTCGGAACAGATGTTGGGACCGTCTTAAAGTAGTTCATTTCTTAAGGAG	1600
DB	433	ValMetTyrLeuGlyThrSerThrGlyProLeuHisLysAlaVal---ValProGln---	450
QY	1601	ACTTGGTATGATTTAGAAGAGTCTCTGCTGGAAGAAATGACAGTTCGCGGAACGACT	1660
DB	451	-----AspSerSerAlaTyrLeuValGluGluIleGlnLeuSerProAspSerGlu	467
QY	1661	GCTATTTTCAGCAATGGAGCTTCCACTAAGCAGCAACAACATATATATTTGGTTCAACGGCT	1720
DB	468	ProValArgAsnLeuGlnLeuAlaProAlaGlnGlyAlaValPheAlaGlyPheSerGly	487
QY	1721	GGGGTTCGCCAGCTCCCTTACCCGCTGTGATATTTACGGGAAACGCTGTGCTGAGTGT	1780
DB	488	GlyIleTrpArgValProArgAlaAsnCysSerValTyr---GluSerCysValAspCys	506
QY	1781	TGCTTCGCCCGAGACCTTACTGTCTTGGAGATGTTCTTCGA-----TGTTCTCGCTAT	1834
DB	507	ValLeuAlaArgAspProHisCysAlaTrpAspProGluSerArgLeuCysSer-----	524
QY	1835	TTTCCACTGCAAGAGAGCGCACAGACGACAAAGATATAAGAAATGGAGCCCATGACT	1894
DB	525	LeuLeuSerGlySerThrLysProTrpLysGlnAspMetGluArgGlyAsnProGluTrp	544
QY	1895	CAGTCTTTCAGACTTACACCATGATAATCACCATGCG-----ArgGlyProMetAlaArgSerProArgAsp	557
DB	545	ValCysThr-----	577
QY	1931	CACAGCTTCGAGAGAGAAATC-----ATCTATGTTGTAGAGATAGTAGCATTGTTTG	1984
DB	558	GlnSerProGlnLeuIleLysGluValLeuThrValProAsnSerIleLeuGluLeu	577


```
QY 1544 ATGTTTATCGGAACAGAGTCTTGGACCGTCTCTTAAAGTAGTTTCAATTCCTTAAGGAGACT 1603
Db 460 TyrTyrValGlyThrAsnLeuGlyArgIleTyrLysIleValGln-----474
QY 1604 TGGTATGATTTAGAAGAGTCTTGGTGGGAAGAAATGACAGATTTTTCGG-----GAACCG 1657
Db 475 TyrTyrArgAsnGlyGluSerLeuSerLysLeuLeuAspIlePheGluValAlaProAsn 494
QY 1658 ACTGCTATTTTCAGCAATGAGCTTTCCACTAAGACGACGACAACTATATATTTGGTTCACAG 1717
Db 495 GluAlaIleGlnValMetGluIleSerGlnThrArgLysSerLeuTyrIleGlyThrAsp 514
QY 1718 GCTGGGTTGCCAGCTCCCTTTACACGGTGTGATATTACGGGAAGCGTGTCTGAG 1777
Db 515 HisArgIleLysGlnIleAspLeuAlaMetCysAsnArgArgTyrAspAsnCysPheArg 534
QY 1778 TGTTCCTCGCCGACACCTTACTGTCTTGGGATGGTTCTGCA-----TGTTCCTCGC 1831
Db 535 Cys-----ValArgAspProTyrCysGlyTyrAspLysGluAlaAsnThrCysArgPro 552
QY 1832 TATTTTCCCACTGCCAAGACGACGACAGACAGACAGATATAAGAAATGGAGACCCACTG 1891
Db 553 Tyr-----GluLeuAspLeuLeuGlnAspValAlaAsn-----563
QY 1892 ACTCACTGTTACAGCTTACACCATGATATCACCATGGCCACAGCCCTGAAGAGAGATC 1951
Db 564 ----GluThrSerAspIleCysAspSerSer-----ValLeuLysLysLysIle 578
QY 1952 ATC-----TATGGTGTAGAGAAATAGTACACATTTTGGATGCGTCCGGAAGTCGCAG 2005
Db 579 ValValThrTyrGly-----GlnSerValHisLeuGlyCysPheValLysIlePro 595
QY 2006 AGAGCGCTG-----GTCTATTGGCAATTCAGAGCGCGGAATGAAGAGCGAATA 2053
Db 596 GluValLeuLysAsnGluGlnValThrTyrHisHisSerLysAspLysGlyArg 614
QY 2054 GAAGAGATCAGAGTGGAT-----GATCATATCATCAGACAGATCAAGGCCCTCTGCTA 2107
Db 615 TyrGluIleArgTyrSerProThrLysIleGluThrThrGluArgGlyLeuValVal 634
QY 2108 CGTAGCTTACACAGAGGATTCAGCGCAATTAACCTCTGCCAT 2149
Db 635 ValSerValAsnGluAlaAspGlyGlyArgTyrAspCysHis 648

RESULT 13
JH0798
fasciclin IV precursor - American bird grasshopper
C:Species: Schistocerca americana (American bird grasshopper)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C:Accession: JH0798
R:Kolodkin, A.; Matthes, D.J.; O'Connor, T.P.; Patel, N.H.; Admon, A.; Bentley, D.; Good
Neuron 9, 831-845, 1992
A:Title: Fasciclin IV: Sequence, expression, and function during growth cone guidance in
A:Reference number: JH0798, MUID:93040225; PMID:1418998
A:Accession: JH0798
A:Molecule type: mRNA
A:Residues: 1-730 <COL>
A:Cross-references: GB:L00709; NID:gl60844; PID:gl60845
A:Experimental source: embryo
C:Comment: This protein plays a role in growth cone guidance in the developing central n
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-730/Product: fasciclin IV #status predicted <EXT>
F:23-627/Domain: extracellular #status predicted <EXT>
F:628-652/Domain: transmembrane #status predicted <TM>
F:653-730/Domain: intracellular #status predicted <INT>
F:44,71,163,267,360,539/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 5,71e-50 Length: 730
Score: 766.50 Matches: 209
Percent Similarity: 47.70% Conservative: 112
Best Local Similarity: 31.05% Mismatches: 225
```

```
Query Match: 15.66% Indels: 127
DB: 2 Gaps: 29
US-09-774-490-1 (1-2709) x JH0798 (1-730)
QY 209 TTAAGTATGATTTGCTGCTTTTCTGGGGAGTATTACTTACAGCAAGACGAACTATCAG 268
Db 5 LeuValAlaValAlaAlaLeuLeuTyr---ValAlaLeuHisAlaAla-----20
QY 269 AATGGGAAGAACAATGTG---CAAAGGCTGAATTTATCTTCAAGAAGAAATGTTGGAATCC 325
Db 21 ---TrpValAsnAspValSerProLysMetTyrValGlnPheGlyGlu-----35
QY 326 AACATGTGATCACTTCAATGGCTTGGCCAAAGCTCCAGTATCATACCTTCTCTTTTG 385
Db 36 GluArgValGlnArgPheLeuGly---AsnGluSerHisLysAspHisPheLysLeu 54
QY 386 GATGAGCAACGAGTAGGCTGTATGTTGGACAAAGGATCACATATTTTTCATTCGACCTG 445
Db 55 GluLysAspHisAsnSerLeuValGlyAlaArgAsnIleValTyrAsnIleSerLeu 74
QY 446 GTTAATATCAAGGATTTT---CAAGAATTGTGTGGCCAGTATCTTACACCAAGA 499
Db 75 ArgAspLeuThrGluPheThrGluGlnArgIleGluTrpHisSerSerGlyAlaHisArg 94
QY 500 GATGAATCAAGTGGCTGGAAAGACATCTGTAAGAAATGTCTTAATTTTCATCAAGTA 559
Db 95 GluLeuCysTyrLeuLysGlyLysSer---GluAspAspCysGlnAsnIleVal 113
QY 560 CTTAAGGCATATAATCAGACTCACTTGTACCCCTGTGGAACGGGGCTTTTTCATCAAT 619
Db 114 LeuAlaLysIleAspAspArgValLeuIleCysGlyThrAsnAlaTyrLysProLeu 133
QY 620 TGCACCTACATGAAATTTGGACATCATCTCTGAGGACAATATTTTAACTGGAGAACTCA 679
Db 134 CysArgHisTyrAlaLeu-----LysAspGlyAspTyrValValGluLysGlu 149
QY 680 CATTTTGAACCGCCCTGGGAAGAGTCCATATGACCCCTAGCTGCTCAGACGATCCCTT 739
Db 150 Tyr-----GluGlyArgGlyLeuCysProPheAspProAspHisAsnSerThrAla 167
QY 740 TTAATAGATGAGAATATATCTCTGGAACCTCAGCTGATTTTATGGGCGAGACTTGTCT 799
Db 168 TyrSerGluGlyGlnLeuTyrSerAlaThrValAlaAspPheSerGlyThrAspProLeu 187
QY 800 ATCTTCCGAATCTTTGGGCACACCCCAATCAGACAGACGACGATGATCCAGGTGG 859
Db 188 IleTyrArg-----GlyProLeuArgThrGluArgSerAspLeuLysGln 202
QY 860 CTCAATGATCCAAAGTTTCATTAGTGCCACCTCATCTCAGAGAGTGACAATCTCTGAAGAT 919
Db 203 LeuAsnAlaProAsnPheValAsnThr-----MetGluTyrAsn 215
QY 920 GACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGGAGAACACTCTGGAAAGCT 979
Db 216 AspPheIlePhePhePheArgGluThrAlaValGluTyrIleAsnCysGlyLysAla 235
QY 980 ACTCAGCTAGATAGGTGAGATATGCAAGATGACTTTGGAGGCCACAGAGTCTGTG 1039
Db 236 IleTyrSerArgValAlaArgValCysLysHisAspLysGlyGlyProHisGlnPheGly 255
QY 1040 AATAAATGGCAACATCTCTCAAGCTCGTCTGATTTTCTCAGTCCAGGTCCAAATGCG 1099
Db 256 AspArgIlePhePhePheLeuLysSerArgLeuAsnCysSerValProGly-----Asp 273
QY 1100 ATTGACATCACTATTTGATGAAGTCCAG-----GATGTATTCTTAATGAACCTTAAA 1150
Db 274 TyrProPheTyrPheAsnGluIleGlnSerThrSerAspIleIleGluGlyAsnTyrGly 293
QY 1151 GATCTTAAATCCAGTTGTATATGGAGTGTTTAGCACTTTTACGATTAACATTTTCAAGGA 1210
Db 294 GlyGlnValGluLysLeuIleTyrGlyValPheThrThrProValAsnSerIleGlyGly 313
```


Db 222 AlaAspThrPheGluGlyGlnPheLysGluGlnThrGlyLeuAsnSerAsnTrpLeuPro 241
QY 1304 TAT---CAAGGAAGAGTCCCTTATCCAGCGCCAGGAATGTGTCGCCAGAAACATTGGT 1360
Db 242 ValAsnAsnAlaLysValProAspProArgProGlySerCys----- 255
QY 1361 GGTITTTGACTCTACAAGACCTTCTGATGATGTTATTAACCTTTCCAGAGTCAATCCA 1420
Db 256 ---HisAsnAspSerArgAlaLeuProAspProThrLeuAsnPheIleLysThrHisSer 274
QY 1421 GCCATGTACAATCCAGTGTCTTATGAACAATCGCCCAATAGTATGATCAAAACCGATGTA 1480
Db 275 LeuMetAspGluAsnValProAlaPheSerGlnProIleLeuValArgThrSerThr 294
QY 1481 AATTATCAATTTACAAATTCCTGAGAC---CGAGTGGATGCAAGAGATGACAG--- 1534
Db 295 IleTyrArgPheThrGlnIleAlaValAspAlaGlnIleLysThrProGlyGlyLysThr 314
QY 1535 TATGATGTTATCTTATCGGAACAGATGTTGGACGGTCTTAAAGTAGTTTCAATTCT 1594
Db 315 TyrAspValIlePheValGlyThrAspHisGlyLysIleIleLysSerValAsnAlaGlu 334
QY 1595 AAGGAGACTTGTATGAT---TTAGAAGAGGTCTGCTGGAAGAATGACAGCTTTTCGG 1651
Db 335 SerAlaAspSerAlaAspLysValThrSerValValIleGluGluLeuAspValLeuThr 354
QY 1652 GAACCGACTGTGTTTCAGCAATGAGCTTCCATTAAGACACA--- 1696
Db 355 LysSerGluProIleArgAsnLeuGluIleValArgThrMetGlnTyrAspGlnProLys 374
QY 1697 -----CAACTATATTTGTTCAACGGCTGGGGTCCCGAGTCT 1735
Db 375 AspGlySerTyrAspAspGlyLysLeuIleValThrAspSerGlnValValAlaIle 394
QY 1736 CTTTACACCGGTGT---GATATTTACGGGAAGCGTGTGCTGAGTGTGCTCCCGCA 1792
Db 395 GlnLeuHisArgCysHisAsnAspLysIleThrSerCysSerGlyCysValAlaLeuGln 414
QY 1793 GACCTTACTGTGCTGGATGGTCTGCA-----TGT----- 1825
Db 415 AspProTyrCysAlaTrpAspLysIleAlaGlyLysCysArgSerHisGlyAlaProArg 434
QY 1826 -----TCTCGCTATTTT-----CCC 1840
Db 435 TrpLeuGluGluAsnTyrPheTyrGlnAsnValAlaThrGlyGlnHisAlaAlaCysPro 454
QY 1841 ACTGCAAGAGACGACAGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1900
Db 455 SerGlyLysIleAsnSerLys-----AspAlaAsnAlaGlyGluGlyLysGlyPheArg 472
QY 1901 TCAGACTTACCATGATATACCATGCGCCACAGCCCTGGAAGAGAGATCATCATGGT 1960
Db 473 AsnAspMetAspLeuLeuAspSerArgArgGlnSerLysAspGlnGluIleIleAspAsn 492
QY 1961 GTAGAGATAGTAGACATTTTGGATCGAGTCCGAGTCCGAGAGCGGCTGTCTAT 2020
Db 493 IleAspLysAsnPheGlu-----GlyProGlnThrSerAlaAspIleIleAsn 508
QY 2021 TGGCAATTCAGAGCGCAATGAAGACGCAAGAGAGAGATCAGATGATCATCATCAT 2080
Db 509 AlaGlnTyr-----ThrValGluThrLeuVal 517
QY 2081 ATCAGACAGATCAAGGCTTCTGCTAGTCTTACACAGAGAGATTCAGGCAATTAC 2140
Db 518 MetAlaValLeuAlaGlySerIlePheSerLeuLeuValGlyPhePheThrGlyThrPhe 537
QY 2141 CTC-----TGCCATCGGTGGAACATGGTTTCATACAACTCTTCTTAAGTAAACC 2191
Db 538 CysGlyArgArgCysHisLysAspGluAspAsp-----Asn 549
QY 2192 CTGGAAGTCAATTCACACA-----GAGCATTTTGAAGAACTTCTTTCATAAAGATGATGAT 2245

550 LeuProTyrProAspThrGluTyrGluTyrPheGluGlnArgGlnAsnValAsnSerPhe 569
2246 GGAGATGCTCTAGACCAAAAGAAATGCTCAATAGCATGACACCTAGCACAGAGTCTGG 2305
Db 570 ProSerSerCysArgIleGlnGlnGluProLysLeuLeuProGlnValGluGluValThr 589
QY 2306 TACAGAGACTTCATGACGCTCATCAACCCACCCCAATCTCAACACGATGATGATGCTGT 2365
Db 590 TyrAlaGluProValLeuLeuProGlnProProProProAsnLysMet----- 605
QY 2366 GAACAAGTTTGGAAAAAGGACCGAAACCAACGTCGGCAAGGCCAGGACATACCCAGG 2425
Db 606 -----HisSerProLys 609
QY 2426 AACAGTACAAATGGAAGCACTTACAGAAATAAGAAAGGTAGNAACAGAGGACC 2482
Db 610 AsnThrLeuArgLysProProMetHisGlnMetHisGlnGlyProAsnSerGluThr 628
RESULT 15
JC5928
semaphorin F precursor - human
C:Species: Homo sapiens (man)
C:Date: 10-Apr-1998 #sequence_revision 08-May-1998 #text_change 17-Nov-2000
C:Accession: JC5928
R:Simmons, A.D.; Puschel, A.W.; McPherson, J.D.; Overhauser, J.; Lovett, M.
Biochem. Biophys. Res. Commun. 242, 685-691, 1998
A:Title: Molecular cloning and mapping of human semaphorin F from the Cri-du-chat candi
A:Reference number: JC5928; MUID:98125554; PMID:9464278
A:Accession: JC5928
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1074 <STM>
A:Cross-references: GB:U52840; NID:g272583; PIDN:AAC09473.1; PID:g272584
A:Experimental source: brain
C:Comment: This protein disrupts normal brain development and leads to some of the featu
C:Genetics:
A:Gene: semaf
C:Superfamily: human semaphorin F; thrombospondin type 1 repeat homology
F:1-20/Domain: signal sequence #status predicted <SIG>
F:50-533/Domain: semaphorin #status predicted <SEM>
F:840-896/Domain: thrombospondin type 1 repeat homology <THR3>
F:971-993/Domain: transmembrane #status predicted <TM>
Alignment Scores:
Pred. No.: 3,598-47 Length: 1074
Score: 730.00 Matches: 185
Percent Similarity: 51.25% Conservative: 102
Best Local Similarity: 33.04% Mismatches: 205
Query Match: 14.91% Indels: 68
Gaps: 22
US-09-774-490-1 (1-2709) x JC5928 (1-1074)
QY 176 AAAGGACCTACAGCGTCTGCAGCATGGCTGGTGTAACTAGGATGTCTGTCTTTCTGG 235
Db 2 LysGlyThr-----CysVallealatrPheSerSerLeuGlyLeuTrpArg 18
QY 236 GGAGTATTACTTACAGCAAGACCAATCATCAAGTGGGAAGACAACTATGTCGCAAGGCTG 295
Db 19 LeuAlaHisProGluAlaGlnGlyThrThrGlnCysGlnArgThrGluHisPro----- 36
QY 296 AAATTATCTCAAGAAGAAATG-----TTGGAATCCCAACAATGTGATC 337
Db 37 VallieserTyrLysGluIleGlyProTrpLeuArgGluPheArgAlaLysAsnAlaala 56
QY 338 ACTTTCAATGCTTGGCCACACAGCTCCAGTTCATACCTTCCTTTTCGATGAGGAACGG 397
Db 57 AspPheSerGlnLeu-----ThrPhe-----AspProGlyGln 67
QY 398 AGTAGGCTGTATGTTGGAGCAAGATGACATATTTTCGACCTGGTTAATATCAAG 457
Db 68 LysGluLeuValGlyAlaArgAsnTyrLeuPheArgLeuGlnLeuGluAspLeuSer 87

Db	414	GluAlaLeuValHisIleIleTyrLeuAlaThrAspTyrGlyThrIleTyrLysValArg	433
Qy	1586	TCAATTCTAAGAGAGACTTCGGTATGATTTTAAAGACAGGTTCTCGTGGGAAGAAATTCACAGATT	1645
Db	434	ValProLeuAsnGlnThr-----SerSerSerCysLeuLeuGluGluIleGluLeu	450
Qy	1646	TTT-----CGGGAACCGACTGCTGCTATTTCAGCAATGAGCTTTCCACTAAGCAG	1693
Db	451	PheProGluArgArgGluPro-----IleArgSerLeuGlnIleLeuHisSerGln	468
Qy	1694	CAACAACATATATTGGTTCAACGGCTGGGGTTGCCAGCTCCCTTTTACACCGGTGTGAT	1753
Db	469	SerValLeuPheValGluLeuArgGluHisValValIleProLeuLysArgCysGln	488
Qy	1754	ATTTACGGGAAACGGTGTGTGCTGAGTGTGCTCCGCCGAGACCTTACTGTCTGTGGGAT	1813
Db	489	PheTyr---ArgThrArgSerThrCysIleGlyValAlaGlnAspProTyrCysGlyTyrPasp	507

Search completed: August 3, 2003, 09:56:58
Job time : 90 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: July 31, 2003, 11:37:35 ; Search time 9579 Seconds
(without alignments)
11569.494 Million cell updates/sec

Title: US-09-774-490-1
Perfect score: 2709
Sequence: 1 aatctttatttcatgatg.....agggttttttctctaatacc 2709

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2709	100.0	2709	6	AX207154	AX207154 Sequence
2	2508.4	92.6	2530	6	AX743092	AX743092 Sequence
3	2508.4	92.6	2530	6	BD084822	BD084822 Diagnosis
4	2508.4	92.6	2530	9	HUMHSEM	L26081 Homo sapien
5	2508.4	92.6	2530	11	G31703	G31703 SWSS1973 Er
6	2508.4	92.6	2601	6	AR040598	AR040598 Sequence
7	2508.4	92.6	2601	6	I47054	I47054 Sequence 53
8	2050	75.7	5952	10	MUSC1	BD5028 Mus musculus
9	2027.6	74.8	2913	10	MNRNASEMD	X85993 M.musculus
10	1909.4	70.5	2319	10	RNSIICNI	X95286 R.norvegicu
11	1883.8	69.5	2319	10	MUSCOSE	L41541 Mus musculus
12	1744	64.4	3263	5	GGU02528	U02528 Gallus gall
13	1612.4	59.5	1998	10	MUSSEMAIII	L40484 Mus musculus
14	1477	54.5	2325	5	AY030051	AY030051 Xenopus 1
15	1415.2	52.2	1481	6	I11900	I11900 Sequence 1
16	1069.2	39.5	2337	5	AF083382	AF083382 Danio rer
17	1035.8	38.2	3148	5	AF086761	AF086761 Danio rer
18	718.6	26.5	2331	6	E27342	E27342 Novel polyp
19	718.6	26.5	3871	6	AX376280	AX376280 Sequence
20	718.6	26.5	3871	6	AX697241	AX697241 Sequence
21	718.6	26.5	3880	6	E27343	E27343 Novel polyp
22	697.2	25.7	2715	5	GGU28240	U28240 Gallus gall
23	652	24.1	169408	2	AC073110	AC073110 Homo sapi
24	650.4	24.0	179640	9	AC006322	AC006322 Homo sapi
25	647.4	23.9	6474	6	AX743098	AX743098 Sequence
26	647.4	23.9	6474	9	AB002329	AB002329 Human mRN
27	615.6	22.7	2615	5	AF022947	AF022947 Gallus ga
28	599.4	22.1	2337	6	AX155172	AX155172 Sequence
29	596.4	22.0	2898	10	AF034744	AF034744 Mus muscu
30	595.8	22.0	2872	10	MNRNASEMA	X85990 M.musculus
31	594.8	22.0	2328	10	MMSEMH	Z80941 M.musculus
32	593.8	21.9	2879	5	AF124485	AF124485 Danio rer
33	581.6	21.5	2919	9	AX743094	AX743094 Sequence
34	581.6	21.5	2919	9	HSU28369	U28369 Homo sapien
35	569.6	21.0	2813	9	AF217991	AF217991 Homo sapi
36	568.8	21.0	2809	9	BC024220	BC024220 Homo sapi
37	568.8	21.0	2825	9	BC013975	BC013975 Homo sapi
38	566.6	20.5	2300	9	AB083186	AB083186 Homo sapi
39	554.4	20.5	2975	6	BD106505	BD106505 Human sem
40	554.4	20.5	4996	9	BC030690	BC030690 Homo sapi
41	553.6	20.4	3988	10	NM293948	Z93948 M.musculus
42	553.6	20.4	4466	10	NM293947	Z93947 M.musculus
43	552.8	20.4	5177	6	AX743096	AX743096 Sequence
44	552.8	20.4	5177	9	AB000220	AB000220 Homo sapi
45	551.2	20.3	2336	5	AF022946	AF022946 Gallus ga

ALIGNMENTS

RESULT 1
AX207154
LOCUS AX207154 2709 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 1 from Patent WO0155455.
ACCESSION AX207154
VERSION AX207154.1 GI:15394943
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1
AUTHORS Jin, S.
TITLE Resistance sequences and uses thereof
JOURNAL Patent: WO 0155455-A 1 02-AUG-2001;

FEATURES		Millennium Pharmaceuticals, Inc. (US) ; Jin, Shengfang (US)	
source	Location/Qualifiers		
	1..2709		
	/organism="Mus musculus"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:10090"		
BASE COUNT	831 a 560 c 620 g 698 t		
ORIGIN			
	Query Match	100.0%;	Score 2709; DB 6; Length 2709;
	Best Local Similarity	100.0%;	Pred. No. 0;
	Matches 2709;	Conservative 0;	Mismatches 0; Indels 0; Gaps 0;
Qy	1	AATCTTTTATTTATCGATGTTAAACAAGCTTAGTAATCGATGCCACGTCGAGGGGTGTCG	60
Db	1	AATCTTTTATTTATCGATGTTAAACAAGCTTAGTAATCGATGCCACGTCGAGGGGTGTCG	60
Qy	61	ACCACGGCTCGGGAGTAGGTTGAGCTCGCCTGTTCTCCCAATGTCAGCCAGTCTATTT	120
Db	61	ACCACGGCTCGGGAGTAGGTTGAGCTCGCCTGTTCTCCCAATGTCAGCCAGTCTATTT	120
Qy	121	CCAGATTGTTTGAACCTTCTCTGGCCGCACAATACAGGAAGGAAGACTAAAGCAGCAAGG	180
Db	121	CCAGATTGTTTGAACCTTCTCTGGCCGCACAATACAGGAAGGAAGACTAAAGCAGCAAGG	180
Qy	181	GACCTACAGCGTCTGACGATGGCTGGTTAACTAGGATTGTCGTCTTTCTGGGGAGT	240
Db	181	GACCTACAGCGTCTGACGATGGCTGGTTAACTAGGATTGTCGTCTTTCTGGGGAGT	240
Qy	241	ATTACTTACACAGAGCAAACTATCAGAAATGGGAAGAACATGTGCCAAGGCTGAAAT	300
Db	241	ATTACTTACACAGAGCAAACTATCAGAAATGGGAAGAACATGTGCCAAGGCTGAAAT	300
Qy	301	ATCCTACAAAGAAATGTTGGAATCCAACAATGTGATCACTTTCAATGGCTTGGCCAAAG	360
Db	301	ATCCTACAAAGAAATGTTGGAATCCAACAATGTGATCACTTTCAATGGCTTGGCCAAAG	360
Qy	361	CTCAGATTATCATACCTTCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAA	420
Db	361	CTCAGATTATCATACCTTCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAA	420
Qy	421	GGATCACATATTTTTCATTCGACCTGTTTAAATCAAGGATTTTCAAAAGATTTGTGTGCC	480
Db	421	GGATCACATATTTTTCATTCGACCTGTTTAAATCAAGGATTTTCAAAAGATTTGTGTGCC	480
Qy	481	AGTATCTTACACGAGAGATGAATGCAAGTGGCTGGAAAAGACATCTCTGAAAAGATG	540
Db	481	AGTATCTTACACGAGAGATGAATGCAAGTGGCTGGAAAAGACATCTCTGAAAAGATG	540
Qy	541	TGCTAAATTTTCATCAAGGTACTTTAAGGCATATATCAGACTCACTTGTAGCGCTGTGAAC	600
Db	541	TGCTAAATTTTCATCAAGGTACTTTAAGGCATATATCAGACTCACTTGTAGCGCTGTGAAC	600
Qy	601	GGGGGCTTTTCATCCAAATTTGCAACCTACATTGAAATTTGGACATCATCTCTGAGGACAATAT	660
Db	601	GGGGGCTTTTCATCCAAATTTGCAACCTACATTGAAATTTGGACATCATCTCTGAGGACAATAT	660
Qy	661	TTTTAAGCTGGAGAACTCACAATTTTGAACCGGCGTGGGAAGAGTCCATATGACCCTTAA	720
Db	661	TTTTAAGCTGGAGAACTCACAATTTTGAACCGGCGTGGGAAGAGTCCATATGACCCTTAA	720
Qy	721	GCTGCTGACAGCATCCCTTTTAAATAGATGGAGAAATATATCTCTGGAACCTGCAGCTGATTT	780
Db	721	GCTGCTGACAGCATCCCTTTTAAATAGATGGAGAAATATATCTCTGGAACCTGCAGCTGATTT	780
Qy	781	TATGGGGCAGAGCTTTTGCTATCTTCCGAACCTCTTGGGCACCAACCCCAATCAGGACAGA	840
Db	781	TATGGGGCAGAGCTTTTGCTATCTTCCGAACCTCTTGGGCACCAACCCCAATCAGGACAGA	840
Qy	841	GCAGCATGATTCAGGTGGCTCAATGATCCAAAGTTCAATTAGTGGCCACCTCATCTCAGA	900
Db	841	GCAGCATGATTCAGGTGGCTCAATGATCCAAAGTTCAATTAGTGGCCACCTCATCTCAGA	900

QY	901	GAGTGACAAATCCTGAAGATGACAAAGTATATCTTTTCTTCCGTGAAATGCAATAGATGG	960
DB	901	GAGTGACAAATCCTGAAGATGACAAAGTATATCTTTTCTTCCGTGAAATGCAATAGATGG	960
QY	961	AGAACACTCTCTGAAAAAGCTACTCAGCTAGAAATAGGTGAGATATGCAAGAATGACTTTGG	1020
DB	961	AGAACACTCTCTGAAAAAGCTACTCAGCTAGAAATAGGTGAGATATGCAAGAATGACTTTGG	1020
QY	1021	AGGGCACAGAGTCTGGTGAATTAATGACACATTTCTCAAAGCTCGTCTCATTTGCTC	1080
DB	1021	AGGGCACAGAGTCTGGTGAATTAATGACACATTTCTCAAAGCTCGTCTCATTTGCTC	1080
QY	1081	AGTGCCAGGTCCAAATGCGCATTTGACACTCAITTTGATGAATGCAAGGATGTATTCCTAAT	1140
DB	1081	AGTGCCAGGTCCAAATGCGCATTTGACACTCAITTTGATGAATGCAAGGATGTATTCCTAAT	1140
QY	1141	GAACTTTTAAAGATCCTAAAAATCCAGTTGTATATGAGAGTGTATTACGACTTCCAGTAACAT	1200
DB	1141	GAACTTTTAAAGATCCTAAAAATCCAGTTGTATATGAGAGTGTATTACGACTTCCAGTAACAT	1200
QY	1201	TTTCAAGGGATCAGCCGTGTGTATATAGCATGATGTGAGAGGGGTGTTCTTGG	1260
DB	1201	TTTCAAGGGATCAGCCGTGTGTATATAGCATGATGTGAGAGGGGTGTTCTTGG	1260
QY	1261	TCCATATGCCACAGGGATGACCCAACTATCAATGGGTGCTTATCAAGGAAGATGCC	1320
DB	1261	TCCATATGCCACAGGGATGACCCAACTATCAATGGGTGCTTATCAAGGAAGATGCC	1320
QY	1321	CTATCCAGGCGCAGGAATTTGTCCAGCAAAACAATTTGGTGGTGTGACTCTTACAAGGA	1380
DB	1321	CTATCCAGGCGCAGGAATTTGTCCAGCAAAACAATTTGGTGGTGTGACTCTTACAAGGA	1380
QY	1381	CTTCTCTGATGATGTTTAACTTTTGAAGAGTCAATCCAGCAATGTAACAATCCAGTGT	1440
DB	1381	CTTCTCTGATGATGTTTAACTTTTGAAGAGTCAATCCAGCAATGTAACAATCCAGTGT	1440
QY	1441	TCCTATGAACAAATCGCCCAATAGTATCAAAACGGATGTAATTAATTAACAATTTACACAAT	1500
DB	1441	TCCTATGAACAAATCGCCCAATAGTATCAAAACGGATGTAATTAATTAACAATTTACACAAT	1500
QY	1501	TGTCGTAGACCGAGTGGATGCAAGATGGAAGATGGAAGATGATGTTATGTTTATCGGAACAGA	1560
DB	1501	TGTCGTAGACCGAGTGGATGCAAGATGGAAGATGGAAGATGATGTTATGTTTATCGGAACAGA	1560
QY	1561	TGTTGGGACCTCTTAAAGTAGTTTCAATTTCTTAAGGAGCTTGGTATGATTTAGAAGA	1620
DB	1561	TGTTGGGACCTCTTAAAGTAGTTTCAATTTCTTAAGGAGCTTGGTATGATTTAGAAGA	1620
QY	1621	GGTTCCTGCTGGAAGAAATGACAGTGTTCGGGAAACCGACTGCTATTTTCAGCAATGGAGCT	1680
DB	1621	GGTTCCTGCTGGAAGAAATGACAGTGTTCGGGAAACCGACTGCTATTTTCAGCAATGGAGCT	1680
QY	1681	TTCCACTAAGCAGCAACAACTATATATGTTTCAACGGCTGGGGTGGCCAGCTCCCTTT	1740
DB	1681	TTCCACTAAGCAGCAACAACTATATATGTTTCAACGGCTGGGGTGGCCAGCTCCCTTT	1740
QY	1741	ACACGGGTGATATTTTACGGGAAAGCGTGTGCTGAGTGTTCCTCCGCGGAGAGCTTAA	1800
DB	1741	ACACGGGTGATATTTTACGGGAAAGCGTGTGCTGAGTGTTCCTCCGCGGAGAGCTTAA	1800
QY	1801	CTGTGCTGGGATGGTCTGCTGATGTTCTGCTATTTTCCCACTGCAAAAGAGACGCAACAG	1860
DB	1801	CTGTGCTGGGATGGTCTGCTGATGTTCTGCTATTTTCCCACTGCAAAAGAGACGCAACAG	1860
QY	1861	ACGACAAGATATAAAGAAATGAGACCCACTGACTCTGTTTCACTGTTTACACTTACCATGATAA	1920
DB	1861	ACGACAAGATATAAAGAAATGAGACCCACTGACTCTGTTTCACTGTTTACACTTACCATGATAA	1920
QY	1921	TCACATGTCGACAGCCCTGAGGAGAGTCAATCTATGCTGTAGAGATAGTAGCACATT	1980
DB	1921	TCACATGTCGACAGCCCTGAGGAGAGTCAATCTATGCTGTAGAGATAGTAGCACATT	1980
QY	1981	TTTGAATGACAGTCCGAAAGTCGACAGAGAGCGCTGCTTATTTGGCAATTTCCAGAGCGGAAA	2040

913 TGAAGATGACAAAGTATACCTTTTCTTCCGTGAAAAATGCAATAGATGGAGAACACCTCTCGG 972
Db |||||
729 TGAAGATGACAAAGTATACCTTTTCTTCCGTGAAAAATGCAATAGATGGAGAACACCTCTCGG 788
Qy |||||
973 AAAAGTACTACCGCTAGAAATAGGTGAGATATGCAAGAAATGATCTTTGGAGGGGACAGAAG 1032
Db |||||
789 AAAAGTACTACCGCTAGAAATAGGTGAGATATGCAAGAAATGATCTTTGGAGGGGACAGAAG 848
Qy |||||
1033 TCTGGTGAATAATGAGACACATCTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCC 1092
Db |||||
849 TCTGGTGAATAATGAGACACATCTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCC 908
Qy |||||
1093 AAATGSCATTTGACACTCATTTTTCATGAACTGCAAGATGATTTCTTAATGAACTTTAAAGA 1152
Db |||||
909 AAATGSCATTTGACACTCATTTTTCATGAACTGCAAGATGATTTCTTAATGAACTTTAAAGA 968
Qy |||||
1153 TCTAAATAATCCAGTTGTATATGAGTGTGTTACGACTTCCAGTAAACATTTTCAAGGATC 1212
Db |||||
969 TCTAAATAATCCAGTTGTATATGAGTGTGTTACGACTTCCAGTAAACATTTTCAAGGATC 1028
Qy |||||
1213 AGCCGTGTATGTATAGATGATGATGAGAGGGTGTTCCTGTTCCATATGCCA 1272
Db |||||
1029 AGCCGTGTATGTATAGATGATGATGAGAGGGTGTTCCTGTTCCATATGCCA 1088
Qy |||||
1273 CAGGGATGGAACCCAACTATCAATGGTGGCTTATCAAGGAAGAGTCCCTATCCAGGCC 1332
Db |||||
1089 CAGGGATGGAACCCAACTATCAATGGTGGCTTATCAAGGAAGAGTCCCTATCCAGGCC 1148
Qy |||||
1333 AGAACTTGTCCAGCAAAACATTTGGTGGTGTGACTCTCAAGAGACCTTCTCTGATGA 1392
Db |||||
1149 AGAACTTGTCCAGCAAAACATTTGGTGGTGTGACTCTCAAGAGACCTTCTCTGATGA 1208
Qy |||||
1393 TGTATAAATCTTGGCAAGTATCCAGCCATGTATCAATCCAGTGTTCCTATGACAA 1452
Db |||||
1209 TGTATAAATCTTGGCAAGTATCCAGCCATGTATCAATCCAGTGTTCCTATGACAA 1268
Qy |||||
1453 TGCCCCAATAGTGATCAAAACCGATGTAAATTTATCAATTTTACAAATTTGTCTGACCG 1512
Db |||||
1269 TGCCCCAATAGTGATCAAAACCGATGTAAATTTATCAATTTTACAAATTTGTCTGACCG 1328
Qy |||||
1513 AGTGGATGCAAGATGGAAGATGATGATTTATCGGAACAGATGTTGGGACCGT 1572
Db |||||
1329 AGTGGATGCAAGATGGAAGATGATGATTTATCGGAACAGATGTTGGGACCGT 1388
Qy |||||
1573 TCTTAAGTAGTTTCAATCTTAAGCAGACTGTTGATGATTTAGAGAGGTTCTGCTGA 1632
Db |||||
1389 TCTTAAGTAGTTTCAATCTTAAGCAGACTGTTGATGATTTAGAGAGGTTCTGCTGA 1448
Qy |||||
1633 AGAAATGACAGATTTTTCGGGAACCGACTGCTATTTTTCAGCAATGGAGCTTTCCACTAAGCA 1692
Db |||||
1449 AGAAATGACAGATTTTTCGGGAACCGACTGCTATTTTTCAGCAATGGAGCTTTCCACTAAGCA 1508
Qy |||||
1693 GCAACAACTATATATTTGGTTTCAACCGCTGGGTGGCCAGCTCCCTTTTACACCGTGTGA 1752
Db |||||
1509 GCAACAACTATATATTTGGTTTCAACCGCTGGGTGGCCAGCTCCCTTTTACACCGTGTGA 1568
Qy |||||
1753 TATTACGGGAAGCGTGTGCTGAGTGTTCCTCGCCCGAGACCTTACTGTCTGGGA 1812
Db |||||
1569 TATTACGGGAAGCGTGTGCTGAGTGTTCCTCGCCCGAGACCTTACTGTCTGGGA 1628
Qy |||||
1813 TGCTTCTGATGTTCTCGCTATTTTCCCACTGCAAGAGACGCAAGACGCAAGATAT 1872
Db |||||
1629 TGCTTCTGATGTTCTCGCTATTTTCCCACTGCAAGAGACGCAAGACGCAAGATAT 1688
Qy |||||
1873 AGAAATGAGACCCCACTGACTCATCTGTTTACAGCTTACACATGATATATCAATGGCCA 1932
Db |||||
1689 AGAAATGAGACCCCACTGACTCATCTGTTTACAGCTTACACATGATATATCAATGGCCA 1748
Qy |||||
1933 CAGCCCTGAAGAGAGATCATCTATGCTGAGAGATAGTACACATTTTTCGAATGCG 1992
Db |||||
1749 CAGCCCTGAAGAGAGATCATCTATGCTGAGAGATAGTACACATTTTTCGAATGCG 1808

Qy 1993 TCCGAAGTCGACAGAGCGCTGGTCTATTGGCAATTCAGAGCGGAAATGAAGAGCGAAA 2052
Db |||||
1809 TCCGAAGTCGACAGAGCGCTGGTCTATTGGCAATTCAGAGCGGAAATGAAGAGCGAAA 1868
Qy 2053 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAGAGCGCTTCTGCTAGCTAG 2112
Db |||||
1869 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAGAGCGCTTCTGCTAGCTAG 1928
Qy 2113 TCTCAACAGAGAGGATTCAGGCAATTCCTCTGCCATCGGTGGAAACATGGGTTCATACA 2172
Db |||||
1929 TCTCAACAGAGAGGATTCAGGCAATTCCTCTGCCATCGGTGGAAACATGGGTTCATACA 1988
Qy 2173 AACTCTTCTTAAGSTAAACCTCGGAAGTCAITGACAAGAGCAATTTGGAGAACTTCTTCCA 2232
Db |||||
1989 AACTCTTCTTAAGSTAAACCTCGGAAGTCAITGACAAGAGCAATTTGGAGAACTTCTTCCA 2048
Qy 2233 TAAAGATGATGAGAGATGGCTCTAAGACCAAGAAATGTCCATAGCATGACACCTAG 2292
Db |||||
2049 TAAAGATGATGAGAGATGGCTCTAAGACCAAGAAATGTCCATAGCATGACACCTAG 2108
Qy 2293 CCAGAAGTCTGGTACAGAGACTTTCATGCAGCTCATCAACACCCCAATCTCAACACGAT 2352
Db |||||
2109 CCAGAAGTCTGGTACAGAGACTTTCATGCAGCTCATCAACACCCCAATCTCAACACGAT 2168
Qy 2353 GGATGAGTCTGTGCAACAAAGTTTGGAAAGGACCGAAACCAACGCTCGGCAAGCCAGG 2412
Db |||||
2169 GGATGAGTCTGTGCAACAAAGTTTGGAAAGGACCGGAAACCAACGCTCGGCAAGCCAGG 2228
Qy 2413 ACATACCCCGGGAACAGTAAACAAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAAA 2472
Db |||||
2229 ACATACCCCGGGAACAGTAAACAAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAAA 2288
Qy 2473 CAGAGGACCCCAACGAAATTTGAGAGGACCCAGAGAGTCTGAGCTGCATTACCTCTAGA 2532
Db |||||
2289 CAGAGGACCCCAACGAAATTTGAGAGGACCCAGAGAGTCTGAGCTGCATTACCTCTAGA 2348
Qy 2533 AACCTCAAAACAGTAGAAACTTGGCTAGACAAATTAACCTGGAACCAATGCAATATACAT 2592
Db |||||
2349 AACCTCAAAACAGTAGAAACTTGGCTAGACAAATTAACCTGGAACCAATGCAATATACAT 2408
Qy 2593 GAACTTTTTCATGGCATTTATGAGATGTTTACAAATGTTGGGAATTCAGCTGAGTTCCA 2652
Db |||||
2409 GAACTTTTTCATGGCATTTATGAGATGTTTACAAATGTTGGGAATTCAGCTGAGTTCCA 2468
Qy 2653 CCAATTATAAATAAATCCATGAGTAACTTTCTTAATAGGCTTTTTTTC 2702
Db |||||
2469 CCAATTATAAATAAATCCATGAGTAACTTTCTTAATAGGCTTTTTTTC 2518

RESULT 3

BD084822 2530 bp DNA linear PAT 27-AUG-2002
Locus BD084822
DEFINITION Diagnosis method and reagents.
ACCESSION BD084822
VERSION BD084822.1 GI:22630432
KEYWORDS JP 2001522241-A/15.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Leeuwen, F.W.V., Grosveld, F.G. and Burbach, J.P.H.
Patent: JP 2001522241-A 15 13-NOV-2001;
Diagnosis method and reagents

JOURNAL
TITLE
ROTTERDAM, UNIVERSITY OF UTRECHT
OS Homo sapiens (human)
PN JP 2001522241-A/15

COMMENT
PD 13-NOV-2001
PF 02-APR-1998 JP 1998542545
PR 10-APR-1997 US 60/043163
PI FREDERIK W VAN LEEUWEN, FRANKLIN G GROSVELD, JOHANNES PETER
HENRI BURBACH

```
PC C12Q1/68,C07K14/47,C12N15/52,C12N9/00,C12N5/10,A61K38/43, PC
A01K67/027,
PC A01K48/00//C07K16/18
CC Strandedness: Double;
CC Topology: Linear;
CC senaphorin-III gene, Genbank accession number L26081 FH Key
CC Location/Qualifiers
FT source 1..2530
FT Location/Qualifiers
   /organism="Homo sapiens (human)".
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
BASE COUNT 786 a 518 c 576 g 650 t
ORIGIN
Query Match 92.6%; Score 2508.4; DB 6; Length 2530;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 193 CTCGACATGGGCTGTTAACTAGGATTGCTGTCTTTCTGGGAGTAGTATTACTTACAGC 252
DB |||||
9 CTCGACATGGGCTGTTAACTAGGATTGCTGTCTTTCTGGGAGTAGTATTACTTACAGC 68
QY 253 AAGAGCAAACTATCAGAATGGGAAGAACAAATGTGCCAAGGCTCAAAATTATCTTACAAAGA 312
DB |||||
69 AAGAGCAAACTATCAGAATGGGAAGAACAAATGTGCCAAGGCTCAAAATTATCTTACAAAGA 128
QY 313 AATGTTGGATCCAACTATGATCACTTTCAATGGCTTGGGCAACAGCTCCAGTTATCA 372
DB |||||
129 AATGTTGGAAATCAACAATGTGATCACTTTCAATGGCTTGGGCAACAGCTCCAGTTATCA 188
QY 373 TACCTTCTCTTTGGATGAGAAACGAGTAGGCTAGGCTGTATGTTGGAGCAAGGATCAATATT 432
DB |||||
189 TACCTTCTCTTTGGATGAGAAACGAGTAGGCTGTATGTTGGAGCAAGGATCAATATT 248
QY 433 TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGTAGTCTTACAC 492
DB |||||
249 TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGTAGTCTTACAC 308
QY 493 CAGAAGAGATGAATCGAAGTGGGCTGGAAAGACATCTCGAAAGATGTGCTAAATTCAT 552
DB |||||
309 CAGAAGAGATGAATCGAAGTGGGCTGGAAAGACATCTCGAAAGATGTGCTAAATTCAT 368
QY 553 CAAGGTACTTAAGGCATATAATCAGACTCAGCTTGTACGCTGTGGAAACGGGGCTTTTCA 612
DB |||||
369 CAAGGTACTTAAGGCATATAATCAGACTCAGCTTGTACGCTGTGGAAACGGGGCTTTTCA 428
QY 613 TCCAAATTTGCACCTACATTGAAATTTGGACATCATCTCGAGGACAATATTTTAAAGCTGGA 672
DB |||||
429 TCCAAATTTGCACCTACATTGAAATTTGGACATCATCTCGAGGACAATATTTTAAAGCTGGA 488
QY 673 GAACTCACAATTTGAAACGGCGGTGGGAAGTGCATATGACCCTAAGCTGCTGACAGC 732
DB |||||
489 GAACTCACAATTTGAAACGGCGGTGGGAAGTGCATATGACCCTAAGCTGCTGACAGC 548
QY 733 ATCCCTTTTAAATAGATGGGAATTAATCTCTGGAATCTGAGCTGATTTTATGGGGCGAGA 792
DB |||||
549 ATCCCTTTTAAATAGATGGGAATTAATCTCTGGAATCTGAGCTGATTTTATGGGGCGAGA 608
QY 793 CTTTGCTATCTTCCGAACTCTTTGGGCAACCAACCAATTCAGGACAGAGCAGCATGATTC 852
DB |||||
609 CTTTGCTATCTTCCGAACTCTTTGGGCAACCAACCAATTCAGGACAGAGCAGCATGATTC 668
QY 853 CAGGTGGCTCAATGATCCAAATTCATTTAGTGCCCACTCTCAATCTCAGAGAGTGAATCC 912
DB |||||
669 CAGGTGGCTCAATGATCCAAATTCATTTAGTGCCCACTCTCAATCTCAGAGAGTGAATCC 728
QY 913 TCAAGATGACAAGTATATCTTTCTCGGTGAAATGCAATAGATGAGAACACTCTGG 972
DB |||||
729 TGAAGATGACAAGTATATCTTTCTTCGGTGAATGCAATAGATGAGAACACTCTGG 788
```

```
QY 973 AAAAGCTACTCAGCGTAGAATAGGTGAGATATGCAAGAAATGACTTTTGGAGGGCACAGAAG 1032
DB |||||
789 AAAAGCTACTCAGCGTAGAATAGGTGAGATATGCAAGAAATGACTTTTGGAGGGCACAGAAG 848
QY 1033 TCTGTGTAATAATGGACAACTTCTCTCAAGCTGCTGTGATTTGCTCAGTGCCAGGTCC 1092
DB |||||
849 TCTGTGTAATAATGGACAACTTCTCTCAAGCTGCTGTGATTTGCTCAGTGCCAGGTCC 908
QY 1093 AAATGGCAATTCACACTCATTTTGTAGTAATGAGGATGTATTCCTAAATGAATTTTAAAGA 1152
DB |||||
909 AAATGGCAATTCACACTCATTTTGTAGTAATGAGGATGTATTCCTAAATGAATTTTAAAGA 968
QY 1153 TCCTAAATAATCCAGTTGTATATGGAGTGTATACGATTCCTCAAGTAAATTTTCAAGGGATC 1212
DB |||||
969 TCCTAAATAATCCAGTTGTATATGGAGTGTATACGATTCCTCAAGTAAATTTTCAAGGGATC 1028
QY 1213 AGCCGTGTGTATGTATAGCATGATGTGAGAGGGTGTCTTGGTCCATATATGCCCA 1272
DB |||||
1029 AGCCGTGTGTATGTATAGCATGATGTGAGAGGGTGTCTTGGTCCATATATGCCCA 1088
QY 1273 CAGGATGGAACCCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTTATCCACGGCC 1332
DB |||||
1089 CAGGATGGAACCCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTTATCCACGGCC 1148
QY 1333 AGGAACTTTGTCCCAGCAAAACATTTTGGTGGTGTGACTCTCAAGAGGACCTTCTGTATGA 1392
DB |||||
1149 AGGAACTTTGTCCCAGCAAAACATTTTGGTGGTGTGACTCTCAAGAGGACCTTCTGTATGA 1208
QY 1393 TGTATTAACCTTTGCAAGAGTCAATCCAGCCATGTACATCAGTGTGTTTCTATGAACAA 1452
DB |||||
1209 TGTATTAACCTTTGCAAGAGTCAATCCAGCCATGTACATCAGTGTGTTTCTATGAACAA 1268
QY 1453 TCGCCCAATAGTGTATCAAAACGGATGTAATTTATCAATTTACAAATTTGTCGTAGACCG 1512
DB |||||
1269 TCGCCCAATAGTGTATCAAAACGGATGTAATTTATCAATTTACAAATTTGTCGTAGACCG 1328
QY 1513 AGTGGATGCAAGATGAGACATGATGTTTATGTTATCGGAACAGATGTTGGACCGT 1572
DB |||||
1329 AGTGGATGCAAGATGAGACATGATGTTTATGTTTATCGGAACAGATGTTGGACCGT 1388
QY 1573 TCTTAAAGTAGTTTCAATTCCTAAGGAGACTGTTGATATGATTTAGAAGGTTCTGCTGGA 1632
DB |||||
1389 TCTTAAAGTAGTTTCAATTCCTAAGGAGACTGTTGATATGATTTAGAAGAGGTTCTGCTGGA 1448
QY 1633 AGAATGACAGTGTTCGGGAACGAGCTGCTATTTTACAGCAATGAGCTTCCACTTAAGCA 1692
DB |||||
1449 AGAATGACAGTGTTCGGGAACGAGCTGCTATTTTACAGCAATGAGGCTTTTCCACTAAGCA 1508
QY 1693 GCAACAACTATATTTGTTTCAACGGCTGGGGTTCGCCAGCTCCCTTTTACACCGGTGTGA 1752
DB |||||
1509 GCAACAACTATATTTGTTTCAACGGCTGGGGTTCGCCAGCTCCCTTTTACACCGGTGTGA 1568
QY 1753 TATTTACGGGAACGGTGTGCTGAGTGTGCTCGCCGAGACCTTACTGTGCTGGGA 1812
DB |||||
1569 TATTTACGGGAACGGTGTGCTGAGTGTGCTCGCCGAGACCTTACTGTGCTGGGA 1628
QY 1813 TGGTTCTGCAATGTTCTCGCTATTTTCCACTGCAAGAGAGCGCACAGACAGATAT 1872
DB |||||
1629 TGGTTCTGCAATGTTCTCGCTATTTTCCACTGCAAGAGAGCGCACAGACAGATAT 1688
QY 1873 AAGAAATGGAGACCACTGACTCTGTTTCAAGCTTACACCATGATTAATCAACATGGCCA 1932
DB |||||
1689 AAGAAATGGAGACCACTGACTCTGTTTCAAGCTTACACCATGATTAATCAACATGGCCA 1748
QY 1933 CAGCCCTGAAGAGAGATTCATTTATGTTGTAGAAATAGTAGCACATTTTGGAAATGCAG 1992
DB |||||
1749 CAGCCCTGAAGAGAGATTCATTTATGTTGTAGAAATAGTAGCACATTTTGGAAATGCAG 1808
QY 1993 TCCCAAGCTCCGAGAGAGCGGTGCTTATTTCCAGTATTCAGAGCGCAATGAGAGCGGAAA 2052
DB |||||
1809 TCCCAAGCTCCGAGAGAGCGGTGCTTATTTGGCAATTCAGAGCGGCAATGAGAGCGGAAA 1868
QY 2053 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCGCTTCTGCTACGTAG 2112
```

Db	1869	AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAG	1928
Qy	2113	TCTTACACAGAGGATTCAGGCAATTAACCTCTGCCATGCGGTGGACATGGGTTACATCA	2172
Db	1929	TCTTACACAGAGGATTCAGGCAATTAACCTCTGCCATGCGGTGGACATGGGTTACATCA	1988
Qy	2173	AACCTCTTCTTAAGGTAACCTGGAAGTCAATGACACAGAGCATTTGGAAGAACTTCTTCA	2232
Db	1989	AACCTCTTCTTAAGGTAACCTGGAAGTCAATGACACAGAGCATTTGGAAGAACTTCTTCA	2048
Qy	2233	TAAAGATGATGATGAGATGGCTTCTAAGACCAAGAAATGTCCAATGACATGACACCTAG	2292
Db	2049	TAAAGATGATGATGAGATGGCTTCTAAGACCAAGAAATGTCCAATGACATGACACCTAG	2108
Qy	2293	CCAGAGGTCGTGACAGAGACTTCATGACGCTCATCAACACCCCAATCTCAACAGAT	2352
Db	2109	CCAGAGGTCGTGACAGAGACTTCATGACGCTCATCAACACCCCAATCTCAACAGAT	2168
Qy	2353	GGATGAGTCTGTGAAACAAGTTTGGAAAGGGACCGAAACCAACGTCGCGCAAGGCCAGG	2412
Db	2169	GGATGAGTCTGTGAAACAAGTTTGGAAAGGGACCGAAACCAACGTCGCGCAAGGCCAGG	2228
Qy	2413	ACATACCCAGGGAACAGTAACAATGGAAGCACTTCAAGAAATTAAGAAAGGTAGAAA	2472
Db	2229	ACATACCCAGGGAACAGTAACAATGGAAGCACTTCAAGAAATTAAGAAAGGTAGAAA	2288
Qy	2473	CAGGAGGCCAGCAATTTGAGAGGCCACCGAGAGTGTCTGAGCTGCAATACCTCTAGA	2532
Db	2289	CAGGAGGCCAGCAATTTGAGAGGCCACCGAGAGTGTCTGAGCTGCAATACCTCTAGA	2348
Qy	2533	AACCTCAACAGTGAAGTCTGCTAGCAATTAACCTGGAAGAAACCAATGCAATATACAT	2592
Db	2349	AACCTCAACAGTGAAGTCTGCTAGCAATTAACCTGGAAGAAACCAATGCAATATACAT	2408
Qy	2593	GAACTTTTTCATGCGCATATGCGATGTTTCAATGCGTGGAAATTCAGCTGAGTTCCA	2652
Db	2409	GAACTTTTTCATGCGCATATGCGATGTTTCAATGCGTGGAAATTCAGCTGAGTTCCA	2468
Qy	2653	CCAAATTAATAATCAATGATGATTAACCTTCTTAATAGGCTTTTTC	2702
Db	2469	CCAAATTAATAATCAATGATGATTAACCTTCTTAATAGGCTTTTTC	2518
RESULT 4			
HUMSEM			
LOCUS	HUMSEM 2530 bp mRNA linear PRI 08-MAY-1995		
DEFINITION	Homo sapiens semaphorin-III (Hsma-I) mRNA, complete cds.		
ACCESSION	L26081		
VERSION	L26081.1 GI:799328		
KEYWORDS	semaphorin		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 2530)		
JOURNAL	The semaphorin genes encode a family of transmembrane and secreted		
MEDLINE	growth cone guidance molecules		
PUBMED	Cell 75 (7), 1389-1399 (1993)		
COMMENT	8269517		
FEATURES	On May 8, 1995 this sequence version replaced gi:436559.		
source	Original source text: Homo sapiens (human).		
Location/Qualifiers	1..2530		
mol_type="mRNA"	/organism="Homo sapiens"		
/db_xref="taxon:9606"	/mol_type="mRNA"		
/tissue_type="brain"	/db_xref="taxon:9606"		
/dev_stage="foetus"	/tissue_type="brain"		
1..2530	/dev_stage="foetus"		
/gene="Hsma-III"	1..2530		
gene	/gene="Hsma-III"		

CDS	786 a	518 c	576 g	650 t
BASE COUNT	786 a	518 c	576 g	650 t
ORIGIN	PGSNKWKHLQENKKGRNRRTHEFERAPRSV"			
Query Match	92.6%; Score 2508.4; DB 9; Length 2530;			
Best Local Similarity	100.0%; Pred. No. 0;			
Matches 2509; Conservative	0; Mismatches 1; Indels 0; Gaps 0;			
Qy	193	CTGCAGCATGGCTGGTTAACTAGGATTTGCTGCTTTTCTGGGAGATTACTTTACAGC	252	
Db	9	CTGCAGCATGGCTGGTTAACTAGGATTTGCTGCTTTTCTGGGAGATTACTTTACAGC	68	
Qy	253	AAGAGCAACTATCAGATGGAAGAAACAATGTGCAAGGCTGAAATTTATCTTACAAAGA	312	
Db	69	AAGAGCAACTATCAGATGGAAGAAACAATGTGCAAGGCTGAAATTTATCTTACAAAGA	128	
Qy	313	AATGTTGGAATCCAACTATGATCATTCAATGGCTTGGCCAAACAGCTCCAGTTATCA	372	
Db	129	AATGTTGGAATCCAACTATGATCATTCAATGGCTTGGCCAAACAGCTCCAGTTATCA	188	
Qy	373	TACCTTCCTTTTGGATGAGGAACGGAGTAGCTGTATGTTGGAGCAAGGATCACATATT	432	
Db	189	TACCTTCCTTTTGGATGAGGAACGGAGTAGCTGTATGTTGGAGCAAGGATCACATATT	248	
Qy	433	TTCATTCACCTGGTTTAAATATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACAC	492	
Db	249	TTCATTCACCTGGTTTAAATATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACAC	308	
Qy	493	CAGAAGAGATGAATGCAAGTGGCTGGAAAAAGACATCTGAAAGAAATGTCTTAATTCAT	552	
Db	309	CAGAAGAGATGAATGCAAGTGGCTGGAAAAAGACATCTGAAAGAAATGTCTTAATTCAT	368	
Qy	553	CAAGGTACTTAAGGCATATATCAGATCATTGTGACGCTGTGGAACGGGGCTTTTCA	612	
Db	369	CAAGGTACTTAAGGCATATATCAGATCATTGTGACGCTGTGGAACGGGGCTTTTCA	428	
Qy	613	TCCAAATTTGACCTACATTTGAAATTTGGACATCATCTGAGGACAAATATTTTAAAGCTGA	672	
Db	429	TCCAAATTTGACCTACATTTGAAATTTGGACATCATCTGAGGACAAATATTTTAAAGCTGA	488	
Qy	673	GAACTCACATTTTGAAAAACGGCCGTGGGAAGAGTCCATATACCCCTAAGCTGCTGACAGC	732	
Db	489	GAACTCACATTTTGAAAAACGGCCGTGGGAAGAGTCCATATACCCCTAAGCTGCTGACAGC	548	
Qy	733	ATCCCTTTTATAGATGAGAAATTTATCTCTGGAACCTGACGCTGATTTTATGGGGCGAGA	792	
Db	549	ATCCCTTTTATAGATGAGAAATTTATCTCTGGAACCTGACGCTGATTTTATGGGGCGAGA	608	
Qy	793	CTTTGCTATCTTCGAACTCTTTGGGCACACACCCCAATCAGGACAGACAGCATGATTC	852	
Db	609	CTTTGCTATCTTCGAACTCTTTGGGCACACACCCCAATCAGGACAGACAGCATGATTC	668	
Qy	853	CAGGTGGCTCAATGATCCAAAGTTTCAATTTAGTGGCCACCTCATCTCAGAGAGTGACAATCC	912	

Db 669 CAGGTGGCTCAATGATCCAAAGTTTCATTAGTGCACCTCATCTCAGAGAGTGACAATCC 728
QY 913 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGGAGAACACTCTGG 972
Db 729 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGGAGAACACTCTGG 788
QY 973 AAAAGCTACTCAGCTAGATAGGTGAGATATGCAAGATGACTTTGGGGGACAGAAG 1032
Db 789 AAAAGCTACTCAGCTAGATAGGTGAGATATGCAAGATGACTTTGGGGGACAGAAG 848
QY 1033 TCTGGTGAATAAATGAGCAACATCTCTCAAGCTCGTCTGATTGCTCAGTSCCAGGTCC 1092
Db 849 TCTGGTGAATAAATGAGCAACATCTCTCAAGCTCGTCTGATTGCTCAGTSCCAGGTCC 908
QY 1093 AAATGGCAATGACACTCAATTTTGGTGAAGTGCAGGATGATTCTTAATGAATCTTTAAAGA 1152
Db 909 AAATGGCAATGACACTCAATTTTGGTGAAGTGCAGGATGATTCTTAATGAATCTTTAAAGA 968
QY 1153 TCCTAAAAATCCAGTTGTATATGAGTGTGTTAGGACTTCCAGTAAATTTCAAGGGATC 1212
Db 969 TCCTAAAAATCCAGTTGTATATGAGTGTGTTAGGACTTCCAGTAAATTTCAAGGGATC 1028
QY 1213 AGCCGTGTGTATGTATAGCATGAGTGTGTTAGGAGGGTGTCTTGGTCCATATGCCCA 1272
Db 1029 AGCCGTGTGTATGTATAGCATGAGTGTGTTAGGAGGGTGTCTTGGTCCATATGCCCA 1088
QY 1273 CAGGGATGGACCAACTATCAATGGGTGCTTATCAAGGAAGTCCCTATCCACGGCC 1332
Db 1089 CAGGGATGGACCAACTATCAATGGGTGCTTATCAAGGAAGTCCCTATCCACGGCC 1148
QY 1333 AGGAATGTGCCAGCAAAACATTTGGTGGTGTGCTCTCAAGAGGACTTCTCTGATGA 1392
Db 1149 AGGAATGTGCCAGCAAAACATTTGGTGGTGTGCTCTCAAGAGGACTTCTCTGATGA 1268
QY 1393 TGTATTAACCTTTGCAAGAGTCAATCCAGGAGTGTGTTAGGAGGGTGTCTCTGATGA 1452
Db 1209 TGTATTAACCTTTGCAAGAGTCAATCCAGGAGTGTGTTAGGAGGGTGTCTCTGATGA 1268
QY 1453 TCGCCCAATGATGATCAAAACGGATGTAATTTATCAAAATTTGCTGAGACCG 1512
Db 1269 TCGCCCAATGATGATCAAAACGGATGTAATTTATCAAAATTTGCTGAGACCG 1328
QY 1513 AGTGATGACAGAAGATGGACATGATGTTATGTTTATCGGAAACAGATGTTGGACCGT 1572
Db 1329 AGTGATGACAGAAGATGGACATGATGTTATGTTTATCGGAAACAGATGTTGGACCGT 1388
QY 1573 TCTTAAAGTAGTTTCAATTCCTAAGGAGTGTGTTAGTATGTTTAGAGAGTGTCTCTGGA 1632
Db 1389 TCTTAAAGTAGTTTCAATTCCTAAGGAGTGTGTTAGTATGTTTAGAGAGTGTCTCTGGA 1448
QY 1633 AGAAATGACAGTGTTCGGGAACCGACTGCTATTTTTCAGCAATGGAGCTTTCACCTAAGCA 1692
Db 1449 AGAAATGACAGTGTTCGGGAACCGACTGCTATTTTTCAGCAATGGAGCTTTCACCTAAGCA 1508
QY 1693 GCAACAACTATATATTTGTTTCAACGGCTGGGGTGGCCAGCTCCCTTTACACCGGTGGA 1752
Db 1509 GCAACAACTATATATTTGTTTCAACGGCTGGGGTGGCCAGCTCCCTTTACACCGGTGGA 1568
QY 1753 TATTTACGGGAAGCGTGTGCTGAGTGTGCTCCCGGAGACCTTACTGTGCTGGGA 1812
Db 1569 TATTTACGGGAAGCGTGTGCTGAGTGTGCTCCCGGAGACCTTACTGTGCTGGGA 1628
QY 1813 TGGTTCTGATGTTCTCGTATTTTCCCACTGCAAGAGACGCAACAGACCAAGATAT 1872
Db 1629 TGGTTCTGATGTTCTCGTATTTTCCCACTGCAAGAGACGCAACAGACCAAGATAT 1688
QY 1873 AAGAAATGAGACCACTGACTGCTGTTTCAAGTGTACACCATGATATACCATGGCCA 1932
Db 1689 AAGAAATGAGACCACTGACTGCTGTTTCAAGTGTACACCATGATATACCATGGCCA 1748
QY 1933 CAGCCCTGAGAGAGAAATCATCTATGTTGTAGAGATAGTACACATTTTGGGAATGAG 1992
Db 1749 CAGCCCTGAGAGAGAAATCATCTATGTTGTAGAGATAGTACACATTTTGGGAATGAG 1808

QY 1993 TCCGAAGTCGAGAGAGCGCTGGTCTATTGGCAATTCAGAGGGCAATGAGAGCGAAA 2052
Db 1809 TCCGAAGTCGAGAGAGCGCTGGTCTATTGGCAATTCAGAGGGCAATGAGAGCGAAA 1868
QY 2053 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCCCTTCTGCTACGTAG 2112
Db 1869 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCCCTTCTGCTACGTAG 1928
QY 2113 TCTCAACAGAGAAGATTCAAGCAATTAACCTCTGCCATCGGTGGAAACATGGTTCATACA 2172
Db 1929 TCTCAACAGAGAAGATTCAAGCAATTAACCTCTGCCATCGGTGGAAACATGGTTCATACA 1988
QY 2173 AACTTCTTAAAGTAAACCTTGAAGTCAITGACACAGAGCATTTGGGAAGAACTTCTTCA 2232
Db 1989 AACTTCTTAAAGTAAACCTTGAAGTCAITGACACAGAGCATTTGGGAAGAACTTCTTCA 2048
QY 2233 TAAAGATGATGATGGAGATGGCTCTTAAGACCAAGAAATGTCATAGCATGACACCTAG 2292
Db 2049 TAAAGATGATGATGGAGATGGCTCTTAAGACCAAGAAATGTCATAGCATGACACCTAG 2108
QY 2293 CCAGAAGTCTGGTACAGAGACTTTCATGAGCTCATCAACCAACCCCAATCTCAACACGAT 2352
Db 2109 CCAGAAGTCTGGTACAGAGACTTTCATGAGCTCATCAACCAACCCCAATCTCAACACGAT 2168
QY 2353 GGATGAGTCTGTGAACAAAGTTTGGAAAGGGACCGAAAAACAACGTCCGCAAGGCCAGG 2412
Db 2169 GGATGAGTCTGTGAACAAAGTTTGGAAAGGGACCGAAAAACAACGTCCGCAAGGCCAGG 2228
QY 2413 ACATACCCCGGGAACAGTAAACAAATGGAAGACCTTACAGAAATATAGAAAGGTAGAAA 2472
Db 2229 ACATACCCCGGGAACAGTAAACAAATGGAAGACCTTACAGAAATATAGAAAGGTAGAAA 2288
QY 2473 CAGGAGGACCCAGCAATTTGAGAGGACCCAGAGAGTGTCTGAGCTGCATTACCTCTAGA 2532
Db 2289 CAGGAGGACCCAGCAATTTGAGAGGACCCAGAGAGTGTCTGAGCTGCATTACCTCTAGA 2348
QY 2533 AACTTCAACAAAGTAGAACTTGCCTAGCAATTAACCTGGAAGAAACAAATGCAATATACAT 2592
Db 2349 AACTTCAACAAAGTAGAACTTGCCTAGCAATTAACCTGGAAGAAACAAATGCAATATACAT 2408
QY 2593 GAACTTTTTCATGGCATTATGTCGATGTTTCAATGCTGGGAATTCAGCTGAGTTCCA 2652
Db 2409 GAACTTTTTCATGGCATTATGTCGATGTTTCAATGCTGGGAATTCAGCTGAGTTCCA 2468
QY 2653 CCAATTATAAATTAATCCATGAGTAACTTTCTTAATAGGCTTTTTCCTCC 2702
Db 2469 CCAATTATAAATTAATCCATGAGTAACTTTCTTAATAGGCTTTTTCCTCC 2518

RESULT 5

G31703
LOCUS G31703
DEFINITION SWS1973 Eric D. Green Homo sapiens STS genomic, sequence tagged site.
ACCESSION G31703
VERSION G31703.1 GI:1916428
KEYWORDS STS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2530)
Bouffard G.G., Iyer L.M., Idol J.R., Braden V.V., Cunningham A.F., Weintraub L.A., Mohr-Fidwell R.M., Peluso D.C., Fulton R.S., Leckie M.P. and Green E.D.
A collection of 1814 human chromosome 7-specific STSs
Genome Res. 7 (1), 59-64 (1997)
PUBMED 9037602
REFERENCE 2 (bases 1 to 2530)
Green E.D.
AUTHORS Human chromosome 7 STSs (1997)
TITLE Human chromosome 7 STSs (1997)

Db	2349	AACTCAAACAGTAGAAGAACTTCGCTAGACAATAAATCTGGAAAAACAAATGCAATATACAT	2400
Qy	2593	GAACCTTTTTTCATGGCAATATATGTGGATGTTTACAATGGTGGGAAATTCAGCTGAGTTCCA	2652
Db	2409	GAACCTTTTTTCATGGCAATATATGTGGATGTTTACAATGGTGGGAAATTCAGCTGAGTTCCA	2468
Qy	2653	CCAATTATAAATTAATCCATGATGAACCTTCCTAATAGGCTTTTTTTTCC	2702
Db	2469	CCAATTATAAATTAATCCATGATGAACCTTCCTAATAGGCTTTTTTTTCC	2518
RESULT 6			
LOCUS	AR040598	2601 bp	DNA
DEFINITION	Sequence 53 from patent US 5807826.	linear	PAT 29-SEP-1999
ACCESSION	AR040598		
VERSION	AR040598.1	GI:5959961	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 2601)		
AUTHORS	Goodman,C.S., Kolodkin,A.L., Matthes,D., Bentley,D.R. and O'Connor,T.		
TITLE	Smaphoxin gene family		
JOURNAL	Patent: US 5807826-A 53 15-SEP-1998;		
FEATURES	Location/Qualifiers		
source	1..2601		
BASE COUNT	809 a 533 c 593 g 666 t		
ORIGIN			
Query Match	92.6%;	Score 2508.4;	DB 6; Length 2601;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 2509;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
Qy	193	CTCAGCATGGGCTGGTTAACTAGGATGTCTCTTTCTTGGGGAGTATTACTTACAGC	252
Db	9	CTGCAGCATGGGCTGGTTAACTAGGATGTCTCTTTCTTGGGGAGTATTACTTACAGC	68
Qy	253	AAGGCNAACTATCAGANTGGGAGAACAAATGTGCCAAGCTGAAATATCTTACAAAG	312
Db	69	ARGAGCAAACTATCAGAAATGGGAGAAACAAATGTGCCAAGCTGAAATATCTTACAAAG	128
Qy	313	AATGTTGGAAATCAAACAATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCA	372
Db	129	AATGTTGGAAATCAAACAATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCA	188
Qy	373	TACCTTTCCTTTTGGATGAGGAAACGAGTAGTGGCTGTATGTGGAGCAAAAGATCACATAT	432
Db	189	TACCTTTCCTTTTGGATGAGGAAACGAGTAGTGGCTGTATGTGGAGCAAAAGATCACATAT	248
Qy	433	TTCAATTCGACCTGGTTAAATATCAAGATTTTCAAAAGATTTGTGGCCAGTATCTTTACAC	492
Db	249	TTCAATTCGACCTGGTTAAATATCAAGATTTTCAAAAGATTTGTGGCCAGTATCTTTACAC	308
Qy	493	CAGAAGAGATGAATGCAAGTGGCTCGGAAAGACATCTCGAAAGATGTGCTAAATTTTCA	552
Db	309	CAGAAGAGATGAATGCAAGTGGCTCGGAAAGACATCTCGAAAGATGTGCTAAATTTTCA	368
Qy	553	CAAGGTACTTTAAGGCATATAATCAGACTCACTTTGTACGCTGTGGAAACGGGGCTTTTCA	612
Db	369	CAAGGTACTTTAAGGCATATAATCAGACTCACTTTGTACGCTGTGGAAACGGGGCTTTTCA	428
Qy	613	TCCAAATTTGCACTTAATTGAAATTTGGACATCATCTGAGGACAAATATTTTAAAGCTGGA	672
Db	429	TCCAAATTTGCACTTAATTGAAATTTGGACATCATCTGAGGACAAATATTTTAAAGCTGGA	488
Qy	673	GAACTCACATTTTGAACCGGCGTGGGAGAGTCCATATGACCTTAAGCTGCTGACAGC	732
Db	489	GAACTCACATTTTGAACCGGCGTGGGAGAGTCCATATGACCTTAAGCTGCTGACAGC	548
Qy	733	ATCCCTTTTAAATAGATGGAGAATTATACCTCTGGAACTGCAGCTGATTTTATGGGGCGAGA	792

JOURNAL Patent: US 5639856-A 53 17-JUN-1997;
FEATURES Location/Qualifiers
source 1. .2601
BASE COUNT 809 a 533 c 593 g 666 t
ORIGIN

Query Match 92.6%; Score 2508.4; DB 6; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 193 CTGACATGGGCTGGTAACTAGGATTCCTGCTCTTTCTGGGAGTATTACTTACAGC 252
DB 9 CTGACATGGGCTGGTAACTAGGATTCCTGCTCTTTCTGGGAGTATTACTTACAGC 68

QY 253 AAGAGCAAACTATCAGAAATGGGAAGAAACAATGTCACAGGCTGAAATTAATCTTACAAAGA 312
DB 69 AAGAGCAAACTATCAGAAATGGGAAGAAACAATGTCACAGGCTGAAATTAATCTTACAAAGA 128

QY 313 AATGTTGGAATCAACAATGTGATCACTTTCAATGGCTTGGCCACAGCTCCAGTTATCA 372
DB 129 AATGTTGGAATCAACAATGTGATCACTTTCAATGGCTTGGCCACAGCTCCAGTTATCA 188

QY 373 TACCTTCTTTGGATGAGGACGGAGTAGGCTGTATGTTGGAGCAAGGATCAATATT 432
DB 189 TACCTTCTTTGGATGAGGACGGAGTAGGCTGTATGTTGGAGCAAGGATCAATATT 248

QY 433 TTCAATTCGACCTGGTTAAATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACAC 492
DB 249 TTCAATTCGACCTGGTTAAATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACAC 308

QY 493 CAGAGAGATGAATCAAGTGGGCTGGAAAGACATCTCTGAAAGAAATGTCTAAATTTCA 552
DB 309 CAGAGAGATGAATCAAGTGGGCTGGAAAGACATCTCTGAAAGAAATGTCTAAATTTCA 368

QY 553 CAAGGTACTTAAAGGATTAATCAAGTGGGCTGGAAAGACATCTCTGAAAGAAATGTCTAA 612
DB 369 CAAGGTACTTAAAGGATTAATCAAGTGGGCTGGAAAGACATCTCTGAAAGAAATGTCTAA 428

QY 613 TCCAAATTCGACCTGATTAATGGAATGGAATCAATCTGAGGACAAATATTTTAAAGCTGA 672
DB 429 TCCAAATTCGACCTGATTAATGGAATGGAATCAATCTGAGGACAAATATTTTAAAGCTGA 488

QY 673 GAACTCAATTTTGAAGACGGCTGGGAAGAGTCCATATGACCTTAAAGCTGCTGACAGC 732
DB 489 GAACTCAATTTTGAAGACGGCTGGGAAGAGTCCATATGACCTTAAAGCTGCTGACAGC 548

QY 733 ATCCCTTTTAAATGAGATGGAATTAATCACTGGAATCTGAGCTGATTTTATGGGGCGAGA 792
DB 549 ATCCCTTTTAAATGAGATGGAATTAATCACTGGAATCTGAGCTGATTTTATGGGGCGAGA 608

QY 793 CTTTGTCTATCTTCCGAATCTTTGGGACCCACCACCAATCAGGACAGAGCAGCATGATTC 852
DB 609 CTTTGTCTATCTTCCGAATCTTTGGGACCCACCACCAATCAGGACAGAGCAGCATGATTC 668

QY 853 CAGGTGGCTCAATGATCCTAAAGTTCATTTAGTCCCACTCTCATCTCAGAGAGTGACAATCC 912
DB 669 CAGGTGGCTCAATGATCCTAAAGTTCATTTAGTCCCACTCTCATCTCAGAGAGTGACAATCC 728

QY 913 TGAAGATGACAAAGTATATCTTTCTTCCGTGAAATGCAATAGATGAGAAACATCTCTGG 972
DB 729 TGAAGATGACAAAGTATATCTTTCTTCCGTGAAATGCAATAGATGAGAAACATCTCTGG 788

QY 973 AAAAGCTACTCAGCTAGATAGGTCAGATATGCAAGATGACTTTGGGGGCGACAGAG 1032
DB 789 AAAAGCTACTCAGCTAGATAGGTCAGATATGCAAGATGACTTTGGGGGCGACAGAG 848

QY 1033 TCTGGTGAATAATGACAACTATCTTCAAGCTCGTCTGATTTGCTCAGTCCAGGTC 1092
DB 849 TCTGGTGAATAATGACAACTATCTTCAAGCTCGTCTGATTTGCTCAGTCCAGGTC 908

QY 1093 AAATGGCATGACATCTTTTGTGAACTGCAAGATGATTTCTTAATGAATCTTTAAAGA 1152
DB 1152 AAATGGCATGACATCTTTTGTGAACTGCAAGATGATTTCTTAATGAATCTTTAAAGA

DB 909 AAATGGCATGACACTCAATTTTGTGAACTGCAAGATGATTTCTTAATGAATCTTTAAAGA 968

QY 1153 TCCTAAAAATCCAGTTGTATATGAGAGTGTTTACGACTTCCAGTAACAATTTTCAAGGATC 1212

DB 969 TCCTAAAAATCCAGTTGTATATGAGAGTGTTTACGACTTCCAGTAACAATTTTCAAGGATC 1028

QY 1213 AGCCGTGTATGTATAGCATGATGTGAGAAAGGTTGTTCTTGGTCCATATATGCCCA 1272

DB 1029 AGCCGTGTATGTATAGCATGATGTGAGAAAGGTTGTTCTTGGTCCATATATGCCCA 1088

QY 1273 CAGGATGAGCCCAACTATCAATGGTCCCTTATCAAGGAAGAGTCCCTATCCAGGCC 1332

DB 1089 CAGGATGAGCCCAACTATCAATGGTCCCTTATCAAGGAAGAGTCCCTATCCAGGCC 1148

QY 1333 AGGAATTTGTCTCCAGCAAAACAATTTGGTGGTTTGTGACTCTACAAAGGACCTTCTCATGA 1392

DB 1149 AGGAATTTGTCTCCAGCAAAACAATTTGGTGGTTTGTGACTCTACAAAGGACCTTCTCATGA 1208

QY 1393 TGTATAAATCTTTTGAAGATCATCCAGCCATGTAATCAATCAAGTGTCTTCTATGAACAA 1452

DB 1209 TGTATAAATCTTTTGAAGATCATCCAGCCATGTAATCAATCAAGTGTCTTCTATGAACAA 1268

QY 1453 TCGCCCAATAGTGTATCAAAACGGATGTAATTAATCAATTTACAAATTTGCTGAGACCG 1512

DB 1269 TCGCCCAATAGTGTATCAAAACGGATGTAATTAATCAATTTACAAATTTGCTGAGACCG 1328

QY 1513 AGTGATGCAAGATGACATGATGATTTATGTTATCGGAAACAGATTTGGGACCGT 1572

DB 1329 AGTGATGCAAGATGACATGATGATTTATGTTATCGGAAACAGATTTGGGACCGT 1388

QY 1573 TCTTAAAGTGTCTTCAATCTTAAAGGACATTTGGTGTATGATTTAGAAAGGTTCTCTGGA 1632

DB 1389 TCTTAAAGTGTCTTCAATCTTAAAGGACATTTGGTGTATGATTTAGAAAGGTTCTCTGGA 1448

QY 1633 AGAATGACATTTTTCGGGAACCGACTGCTATTTACAGCAATGAGCTTCCACTAAGCA 1692

DB 1449 AGAATGACATTTTTCGGGAACCGACTGCTATTTACAGCAATGAGCTTCCACTAAGCA 1508

QY 1693 GCAACAACTATATATTTGGTTCAACCGCTGGGGTTGGCCAGCTTCCCTTTACACCGGTGGA 1752

DB 1509 GCAACAACTATATATTTGGTTCAACCGCTGGGGTTGGCCAGCTTCCCTTTACACCGGTGGA 1568

QY 1753 TATTTACGGGAAAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGGA 1812

DB 1569 TATTTACGGGAAAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGGA 1628

QY 1813 TGGTCTGGA 1872

DB 1629 TGGTCTGGA 1688

QY 1873 AAGAAATGGACACCCACTGACTCACTGTTTACAGACTTACACCATGATATACCATGGCCA 1932

DB 1689 AAGAAATGGACACCCACTGACTCACTGTTTACAGACTTACACCATGATATACCATGGCCA 1748

QY 1933 CAGCCCTGAAGAGAGAAATCATCTATGTTGTAGAAATAGTAGCAATTTTGGAAATGCGAG 1992

DB 1749 CAGCCCTGAAGAGAGAAATCATCTATGTTGTAGAAATAGTAGCAATTTTGGAAATGCGAG 1808

QY 1993 TCCGAAGTCCGAGAGAGCGCTGCTTATTGCAATTTCCAGAGCGGAAATAGAGCGGAAA 2052

DB 1809 TCCGAAGTCCGAGAGAGCGCTGCTTATTGCAATTTCCAGAGCGGAAATAGAGCGGAAA 1868

QY 2053 AAGAGATGACAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAG 2112

DB 1869 AAGAGATGACAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAG 1928

QY 2113 TCTTAAACAGAGAGATTCAGGCAATTTACCTCTGCCATGCGGTGGAACATGGGTTTCAACA 2172

DB 1929 TCTTAAACAGAGAGATTCAGGCAATTTACCTCTGCCATGCGGTGGAACATGGGTTTCAACA 1988

QY 2173 AACTCTTCTTAAAGTAACTTCCCTGGAAGTCAATTTGACAGAGATTTGGAGAACTTCTTCA 2232

DB 1989 AACTCTTCTTAAAGTAACTTCCCTGGAAGTCAATTTGACAGAGATTTGGAGAACTTCTTCA 2048

```
QY 2233 TAAAGATGATGATGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACTAG 2292
|||
Db 2049 TAAAGATGATGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACTAG 2108
|||
QY 2293 CCAAGAGGCTGGTACAGAGACTTCATGAGCTCATCAACCCCAATCTCAACAGCAT 2352
|||
Db 2109 CCAAGAGGCTGGTACAGAGACTTCATGAGCTCATCAACCCCAATCTCAACAGCAT 2168
|||
QY 2353 GGATGAGTCTGTGACAAAGTTTGGAAAGGACCGAACAACACGTCGCGCAAGCCAGG 2412
|||
Db 2169 GGATGAGTCTGTGACAAAGTTTGGAAAGGACCGAACAACACGTCGCGCAAGCCAGG 2228
|||
QY 2413 ACATACCCAGGGAACAGTAAACAATGGAAGCACTTACAGAAATAAAGAAAGTAGAAA 2472
|||
Db 2229 ACATACCCAGGGAACAGTAAACAATGGAAGCACTTACAGAAATAAAGAAAGTAGAAA 2288
|||
QY 2473 CAGGAGGACCCAGAAATTTGAGAGGACCCAGGAGTGTCTGAGCTGCATTAACCTCTAGA 2532
|||
Db 2289 CAGGAGGACCCAGAAATTTGAGAGGACCCAGGAGTGTCTGAGCTGCATTAACCTCTAGA 2348
|||
QY 2533 AACCTCAACAGTAGAAACTTGCCTAGACAACTAAGTGAAGAAACAAATGCAATATACAT 2592
|||
Db 2349 AACCTCAACAGTAGAAACTTGCCTAGACAACTAAGTGAAGAAACAAATGCAATATACAT 2408
|||
QY 2593 GAACTTTTTCATGGCAATATGTGATGTTTCAATGGTGGGAAATTCAGCTGAGTTCCA 2652
|||
Db 2409 GAACTTTTTCATGGCAATATGTGATGTTTCAATGGTGGGAAATTCAGCTGAGTTCCA 2468
|||
QY 2653 CCAATTATAAATAAATCAATGAGTAACTTCTTAATAGGCTTTTTTTTCC 2702
|||
Db 2469 CCAATTATAAATAAATCAATGAGTAACTTCTTAATAGGCTTTTTTTTCC 2518
|||
```

RESULT 8

```
MUSC1
LOCUS Mus musculus sema3A mRNA for semaphorin 3A, complete cds.
DEFINITION D85028
ACCESSION D85028.2 GI:18389545
VERSION 1 (bases 1 to 2356)
KEYWORDS collapse-1.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 2356)
AUTHORS Taniguchi,M., Yusa,S., Fujisawa,H., Naruse,I., Saga,S., Mishina,M.
and Yagi,T.
TITLE Disruption of semaphorin III/D gene causes severe abnormality in peripheral nerve projection
JOURNAL Neuron 19 (3), 519-530 (1997)
MEDLINE 97470885
PUBMED 9331345
REFERENCE 2 (bases 1 to 5952)
AUTHORS Taniguchi,M.
JOURNAL Direct Submission
TITLE Submitted (02-MAY-1996) Masahiko Taniguchi, Department of Biochemistry and Molecular Biology, Graduate School of Medicine, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, Japan (E-mail:taniguti@m.u-tokyo.ac.jp, Tel:81-3-5802-2925, Fax:81-3-3813-8732)
COMMENT On Jan 28, 2002 this sequence version replaced gi:1313903.
FEATURES
Source
1..5952
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
gene
1..5952
/gene="sema3A"
CDS
652..2970
/gene="sema3A"
/codon_start=1
/product="semaphorin 3A"
```

```
/protein_id="BAA19773.1"
/db_xref="GI:18389545"
/translation="MGWFTGIACLFVGLLTARANYANGKNVPRBLKSLYKEMLSNN
VITNGLANSSSYHTFLLDEERSLYVCAKDHPISFNLVNFKDFKIVPVSVYTRDE
KFWAGKDLKSCANPIKLEAYNQTHLYACGTGAFPHPICTIIEVGHPEDNIFKLQDS
CFNENRGKSPYDKLLTASLIDGELYSGTAFMDRDFAIPTRTGLHHHPNIRTEQDHS
RWLNDPRFISAHLIPESDNPDQKVFYFFRENAIDGHSKATHARIQOICKNDFGGH
RSLVKNWTTTFLKARLICSVPNGIDTFHDELQDVLNMSKDPKNPIVGVFTTSSNI
FKGSACVMSMSDVRVRPLGVAHRDGNVQVYQGRVPVPRPCTPSKTFGGFDSF
KDLDDVTTPARSHPMYNPVPIINRPIKTDVNYQTOIVDVRDAEDQOYDMF
IGTDVTVLVKVSVPKETHDLEELVEEIMVTFREPTTISAMELSTKQQLYIGSTAG
VAQPLHRCDIYKACACELAPDYCAWDGSSCSRYPTAKRRTTRDRDNGDPLTH
CSDLQHDHNRHPSLEERIIYVENSTSTFLCSPKSRALVYQQRNEDRIRH
GDHLIRTEQGLLSLOKDGNYLCHAVEHGMOTLLKVLTEVDITHELELLHKDD
DGDGSKLEKSSMTSPKQVWRDPMOLINHPNLNMTDFECEQVWKRDRKORRQPGH
SQSSNWKHWQESKKGRRTHFERAPRSV"
BASE COUNT 1829 a 1203 c 1214 g 1706 t
Query Match 75.7%; Score 2050; DB 10; Length 5952;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 2281; Conservative 0; Mismatches 295; Indels 14; Gaps 4;
QY 118 TTTCCAGATTGTTGAACTTCTCTGGCCGCAATACAGGAAGGAAGCTGCCGAGCTC 629
|||
Db 570 TTTCCAGATTGTTGAACTTCTCTGGCCGCAATACAGGAAGGAAGCTGCCGAGCTC 629
|||
QY 178 AGGGACCTACAGCTCTGCAGCAGTGGCTGTAACTAGGATGTCTCTCTTTCTGGG 237
|||
Db 630 AGGGACCTCCAGCGTCTGCAGCAGTGGCTGTAACTAGGATGTCTCTCTTTCTGGG 689
|||
QY 238 AGTATTACTTACAGCAAGCAAACTATCAGAAAGGAAGAACTATGCGCAAGCTGAA 297
|||
Db 690 TGTATTACTTACAGCAGCAAACTATGCGCAAGGAAGAACTATGCGCAAGCTGAA 749
|||
QY 298 ATTATCTTACAAAGAAATGTTGGAATCAACATGTGTATCATCTTCAATGGCTTGGCAA 357
|||
Db 750 ATTATCTTACAAAGAAATGTTGGAATCAACATGTGTATCATCTTCAATGGCTTGGCAA 809
|||
QY 358 CAGCTCCAGTTATCATACCTTCTTTTGGATGAGGAAGGAGTAGGTGTATGTTGGAGC 417
|||
Db 810 CAGCTCCAGTTATCATACCTTCTTTTGGATGAGGAAGGAGTAGGTGTATGTTGGAGC 869
|||
QY 418 AAAGATCACATATTTTTCATTCGACCTCTGTTTAAATCAAGGATTTCAAAAGATGTGTG 477
|||
Db 870 AAAAGATCATATATTTTTCATTCGAGTGTGAAATTAAGATTTTCAAAAGATGTGTG 929
|||
QY 478 GCCAGTATCTTACACCAAGAGATGAATCAAGTGGGCTGGAAAGACATCTCTGAAAGA 537
|||
Db 930 GCCAGTATCTTACACCAAGAGATGAATCAAGTGGGCTGGAAAGACATCTCTGAAAGA 989
|||
QY 538 ATGTGCTAATTTTCATCAAGGTACTTAAGGCATATAATCAGACTCACTTGTAGCCTGTGG 597
|||
Db 990 ATGTGCTAATTTTCATCAAGGTACTTAAGGCATATAATCAGACTCACTTGTAGCCTGTGG 1049
|||
QY 598 AACGGGGCTTTTTCATCCAAATTTGACCTATCAATGAAATGGACATCATCTCTGAGGACAA 657
|||
Db 1050 AACTGGGGCTTTTTCATCCAAATTTGACCTATCAATGAAATGGACATCATCTCTGAGGACAA 1109
|||
QY 658 TATTTTAAAGCTGGAGAACTCAATTTTGAAGAACGGCGCTGGGAAGAGTCCATATGACCC 717
|||
Db 1110 CATTTTAAAGCTGGAGAACTCAATTTTGAAGAACGGCGCTGGGAAGAGTCCATATGATCC 1169
|||
QY 718 TAAGCTGCTGACAGCATCCCTTTTAAATAGATGGAATATATCTCTGAACTGGAGCTGA 777
|||
Db 1170 CAAACTACTGACTGCTCTCTTCTTAAATAGATGGAATATATCTCTGAACTGGAGCTGCGGA 1229
|||
QY 778 TTTTATGGGGGAGACTTTTGTCTATCTTCCGAATCTTTGGGGCACCACCAACCAATCAGGAC 837
|||
Db 1230 CTTTATGGGGGAGACTTTGTCTATCTTCCGAATCTTTGGGGCACCACCAATCAGGAC 1289
|||
QY 838 AGACGACATGATTTCCAGGTGGCTCAATGATCCAAAGTTTATAGTGGCCACCTCATCTC 897
|||
```

Db 1290 GGAGCAGATGACTCCCGTGGCTCAATGATCTAGATTCATCATGTCGACCATCTCATCCC 1349
QY 898 AGAGAGTGCAATCTCTGAAGATGACAAAGTATATCTTTCTTCCGTGAAATGCAATAGA 957
Db 1350 AGAGAGTGCAATCTCTGAAGATGACAAAGTATATCTTTCTTCCGTGAAATGCAATAGA 1409
QY 958 TGAGAGCACTCTGGAAGAGTACTCACCGCTAGATAGTATAGTATGCAAGATGACTT 1017
Db 1410 CGGAGAACACTCTGGAAGAGCACTCATGCTAGATAGTATAGTATGCAAGATGACTT 1469
QY 1018 TGAGGGGACAGAGTCTGCTGAATTAATGGCAACATCTCTCAAGCTCGTCTGATTG 1077
Db 1470 TGGTGGACACAGAGTCTTGTGAATAATGGCAACATCTCTCAAGCTCGTCTGATTG 1529
QY 1078 CTCAGTGCCAGGTCCAAATGGCAATGACACTCATTTTGAATGAATGACAGATGATTTCT 1137
Db 1530 CTCGTGCGCGTCCCAATGGCAATGACACTCATTTTGAATGAATGACAGATGATTTCT 1589
QY 1138 AATGAATCTTAAGATCTTAATAATCCAGTGTGTATATGAGTGTGTTACGACTTCCAGTAA 1197
Db 1590 AATGAATCTTAAGATCTTAATAATCCAGTGTGTATATGAGTGTGTTACGACTTCCAGTAA 1649
QY 1198 CATTTTCAAGGATCAGCGCTGTGTATGATAGCATGAGTGTGAGAGGGTGTCT 1257
Db 1650 CATCTTTAAGGATCTGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1709
QY 1258 TGGTCCATATGCGCCAGGATGGCCAACTATCAATGGGTGCTTATCAAGGAAGAGT 1317
Db 1710 TGGTCCATATGCTCAGAGATGTGCTTCAACTATCAGTGGGTGCTTATCAAGGAAGAGT 1769
QY 1318 CCCCTATCCAGCGCAGGAATCTGTCAGCAAAACATTTGTTGGTGTGTTGACTTACAAA 1377
Db 1770 CCCCTATCCAGCGCAGGAATCTGTCAGCAAAACATTTGTTGGTGTGTTGACTTACAAA 1829
QY 1378 GGCCTCTCTGATGTATTAACCTTTCAGAGATGATCAGCCATGATCAATCCAGT 1437
Db 1830 GGCCTCTCTGATGTATTAACCTTTCAGAGATGATCAGCCATGATCAATCCAGT 1889
QY 1438 GTTCTCTATGAACATCGCCCAATAGTATCAAAACGGATGTAATTTATCAATTTACACA 1497
Db 1890 GTTCTCTATGAATCGCCCAATAGTATCAAAACGGATGTAATTTATCAATTTACACA 1949
QY 1498 AATGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1557
Db 1950 AATGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2009
QY 1558 AGATGCTGGAGCGGTCTTAAGATGATTTCAATTCCTAAGGAGACTTGTATGATTTAGA 1617
Db 2010 AGATGCTGGAGCGGTCTTAAGATGATTTCAATTCCTAAGGAGACTTGTATGATTTAGA 2069
QY 1618 AGAGGTTCTGCTGGAGAAATGACAGTCTTTCGGGAACCGACTGCTATTTTCAAGCAATGA 1677
Db 2070 AGAGGTTCTGCTGGAGAAATGACAGTCTTTCGGGAACCGACTGCTATTTTCAAGCAATGA 2129
QY 1678 GTTTCCTCAATAGAGCAACAACTATATATGTTTCAACGGCTGGGGTGGCCAGCTCCC 1737
Db 2130 GCTTTCTACTAAACAGCAACAGCTGTATATGGCTCAACTGCGGAGTGGCAGACGCTCC 2189
QY 1738 TTTACACCGGTGTATTTACGGGAAGCGTGTGCTGATGTTGCTGCGCCCGAGACCC 1797
Db 2190 TTTACACCGGTGTATTTACGGGAAGCGTGTGCTGATGTTGCTGCGCCCGAGACCC 2249
QY 1798 TTTACTGTGCTGGAGTGTGCTGATGTTCTGCTATTTTCCACTTGGCAAGAGAGCGAC 1857
Db 2250 TTTACTGTGCTGGAGTGTGCTGATGTTCTGCTATTTTCCACTTGGCAAGAGAGCGAC 2309
QY 1858 AAGACGACAAGATATAAGAAATGGAGACCCACTGACTACTGTTTCAAGCTT---ACACCA 1914
Db 2310 AAGACGACAAGATATAAGAAATGGAGACCCACTGACTACTGTTTCAAGCTT---ACACCA 2369
QY 1915 TGATATACCATGCGCCAGCCCTGAGAGAGATATCATCTATGTTGTATGAGATATGATAG 1974
Db 2370 TGATATATCATGCGCCCGCCCTTGAAGAGAGATATCATCTATGTTGTATGAGTGGAAACAGTAG 2429

QY 1975 CACATTTTGGATGCACTCCGAAAGTCGACAGAGCGCTGCTCTATTGGCAATTCAGAG 2034
Db 2430 TACATTTCTTGGATGCACTCCGAAAGTCGACAGAGCGCTTGGTATATTGGCAATTCAGAG 2489
QY 2035 GCGAAATGAAGAGCGAAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCA 2094
Db 2490 GAGAAATGAAGATCGAAAGAGGAGATCAGAAATGGGTGATCATATCATCAGGACAGAA 2549
QY 2095 AGGCCTTCTGCTACCTAGTCTACACAGAGATTCAGGCAATACCTCTGCGCATGCCGT 2154
Db 2550 AGGCCTTCTGCTACCTAGTCTACACAGAGATTCAGGCAATACCTCTGCGCATGCCGT 2609
QY 2155 GGAACATCGGTTCTATACAAACTCTTCTTAAGGTAAACCTCGAAGTCAITTTGACACAGCA 2214
Db 2610 GGAACATCGGTTCTATACAAACTCTTCTTAAGGTAAACCTCGAAGTCAITTTGACACAGCA 2669
QY 2215 TTTGGAAGAACTCTTCTATAAAGATGATGAGATGGCTCTTAAGCAAAAGAAATGTC 2274
Db 2670 TTTGGAAGAACTCTTCTATAAAGATGATGAGATGGCTCTTAAGCAAAAGAAATGTC 2729
QY 2275 CAATAGCATGACACTAGCCAGAGAGTCTGTCAGAGACTTTCATGACAGCTCATCAACA 2334
Db 2730 GAGCAGCATGACAGCCAGCAGAAAGTCTGTCAGAGACTTTCATGACAGCTCATCAACA 2789
QY 2335 CCCCAATCTCAACAGCATGAGTCTGTCAGCAACAGTCTTGGAAAGGGACCCGAAACA 2394
Db 2790 CCCCAATCTCAACAGCATGAGTCTGTCAGCAACAGTCTTGGAAAGGGACCCGAAACA 2849
QY 2395 ACCTGCGCAAGGCGCAGGACATACCCAGGGAACAGTAAACAAATGGAAGAGCTTACAGA 2454
Db 2850 ACCTGCGCAAGGCGCAGGACATACCCAGGGAACAGTAAACAAATGGAAGAGCTTACAGA 2909
QY 2455 AAATGAAGAGGTAGAACAGGAGGACCCAGAAATTTGAGAGGGCACCAGAGGTCTG 2514
Db 2910 GAGCAAGAGGTAGAACAGGAGGACCCAGAAATTTGAGAGGGCACCAGAGGTCTG 2969
QY 2515 AGCTGCTATCTCTAGAAACCTCAACAGTAGAAATCTGCTAGACAAATAAAGCTGGAAA 2574
Db 2970 AGCTGCTATCTCTAGAAACCTCAACAGTAGAAATCTGCTAGACAAATAAAGCTGGAAA 3029
QY 2575 AACAAATGCAATATACATGAACTTTTTCATGCGGATGATGATGATGATGATGATGATG 2634
Db 3030 ---AAATGCAATATACATGAACTTTTTCATGCGGATGATGATGATGATGATGATGATG 3084
QY 2635 AAATTCAGTCTGAGTTCCACCAATTAATAAATTAATTAATTAATTAATTAATTAATTAAT 2694
Db 3085 AAGTTCAACCGGTTCCACCAATTAATAAATTAATTAATTAATTAATTAATTAATTAATTAAT 3138
QY 2695 TTTTTCCTTA 2704
Db 3139 TTTTTCAGTA 3148

RESULT 9

MMRNASEMD
LOCUS M.musculus mRNA for semaphorin D. 2913 bp mRNA linear ROD 08-JUL-1996
DEFINITION X85993
ACCESSION X85993.1 GI:854329
VERSION semaphorin; semD gene.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
AUTHORS Puschel, A.W., Adams, R.H. and Betz, H.
TITLE Murine semaphorin D/collapsin is a member of a diverse gene family and creates domains inhibitory for axonal extension
JOURNAL Neuron 14 (5), 941-948 (1995)
MEDLINE 95267431
PUBMED 7748561
REFERENCE 2 (bases 1 to 2913)

AUTHORS

Adams, R. H.

TITLE

Direct Submission

JOURNAL

Submitted (30-MAR-1995) R.H. Adams, Max-Planck-Institute fuer

COMMENT

Hirnforschung, Deutscherstr. 46, D- 60528 Frankfurt, FRG

FEATURES

Sequence overlapping with those under the acc#002528, L26080,

source

L26081 & L26082.

FEATURES

Location/Qualifiers

source

1. .2913

gene

/organism="Mus musculus"

CDS

/mol_type="mRNA"

CDS

/strain="NMRI"

CDS

/db_xref="taxon:10090"

CDS

/clone="sema9"

CDS

/clone_lib="Stratagene lambda ZAP"

CDS

/dev_stage="12 days embryo"

CDS

1. .2913

CDS

/gene="semd"

CDS

111. .2429

CDS

/gene="semd"

CDS

/codon_start=1

CDS

/product="sema9 protein D"

CDS

/protein_id="CAA59985.1"

CDS

/db_xref="GI:854330"

CDS

/db_xref="MGI:107558"

CDS

/db_xref="SPTREMBL:Q62180"

CDS

/translation="MGWFTGIACLFWGLLTARANYANGKNVPRLLKLSYKEMLESNN

CDS

VITFNLANSVHYTFLLDESRSLYVGAQKHI FSNLVNI KDFOKIVPVSYTRDE

CDS

CKWAGKDLKECANIKVLEAYNQTHLYACGTGAPHPICTYIEVGHHPEDNIFKLQDS

CDS

HPENGKGSYDPLKLTSLIDGELYSGTAADFWRDPAIFRTLDHHPHIEQDHS

CDS

RWLNDFRSLHLEIIPESNDDEDDKVFYFFRENAIDGHSKATHARIQIICNDPFGH

CDS

RSLNVKWTFLKARLICSVPNGIDTHPELDQVFLMNSKDPKPIVGVFTTSSNI

CDS

FKGSVAVCMYSMSDVRVPLGPAHVDGPNYQVPRPCVPSKPTFGGPDST

CDS

KDLPDVITFGRSHPMYNPVFINNRPIIMIKTDVNYQTOI VDRDVAEDQDYMZ

CDS

IGTDVTLKVSVPFVWHDELEVLEEMVIFREPTTISAMELTKQQLYIGTAG

CDS

VAQLPLHRCDIYKCAECCLARDPYCAWDGSSRYPTAKRTRRQDIRNGDPLRM

CDS

CSDLHDNHGFLSEERI IYGVENSSTPLECSKQALVYVQQRNRRSKRIRM

CDS

GDHIIRTEGGLLRSLQKSDHCHAVEHFMOTLLKVLTEVIDTSHLELLHKDD

CDS

DDGSKI KEMSSMTPSQKVYRDFMOLINHPNLNMTDFEFCBQVWKRDRKQRPPGH

CDS

SOGSNKWKHMDESKKGRNRRTHEPRPSV"

CDS

BASE COUNT 888 a 632 c 679 g 714 t

CDS

ORIGIN

CDS

Query Match

CDS

Best Local Similarity 74.8%; Score 2027.6; DB 10; Length 2913;

CDS

Matches 2267; Conservative 0; Mismatches 309; Indels 14; Gaps 4;

CDS

118 TTTCCAGATGTTTGAACCTTCTCTGCGCCGCAACAATACAGGAAGGAGAAAGCAAGCAAA 177

CDS

29 TTTCCAGATGTTTGAACCTTCTCTGCGCCGCAACAATACAGGAAGGAGAAAGCAAGCTC 88

CDS

178 AGGACCTACAGCTCTGAGCATGGCTGTTTAACTAGGATGCTGCTTTTCTGGG 237

CDS

89 AGGACCTACAGCTCTGAGCATGGCTGTTTAACTAGGATGCTGCTTTTCTGGG 148

CDS

238 AGTATTACTTACAGCAAGAGCAAACTATCAGAAATGGGAAGAAACAATGTGCCAAGGCTGAA 297

CDS

149 TGTATTACTTACAGCAAGAGCAAACTATCAGAAATGGGAAGAAACAATGTGCCAAGGCTGAA 208

CDS

298 ATTATCTTACAAAGAAATGTTGGAATCCAAATGTGATCACTTCAATGGCTTGGCCAA 357

CDS

209 ATTATCGTACAAAGAAATGTTGGAATCCAAATGTGATCACTTCAATGGCTTGGCCAA 268

CDS

358 CAGCTCCAGTATCATACCTCTTTTGGATGAGGACGAGTGGCTGTATGTTGGAC 417

CDS

269 CAGCTCCAGTATCATACCTCTTTTGGATGAGGACGAGTGGCTGTATGTTGGAC 328

CDS

418 AAAGATCATATATTTTCACTCGACCTGTTTAAATATCAAGGATTTTCAAAAGATTTGTG 477

CDS

329 AAAGATCATATATTTTCACTCGACCTGTTTAAATATCAAGGATTTTCAAAAGATTTGTG 388

CDS

478 GCCAGTATCTTACACAGAGAGATGATGCAAGTGGGCTGGGAGGATGTTTGAAGA 537

CDS

389 GCCAGTATCTTACACAGAGAGATGATGCAAGTGGGCTGGGAGGATGTTTGAAGA 448

538 ATGTCGTAATTTTCATCAAGGTACTTTAAGGCATATAATCAGACTCACTTGTAGCCTGTGG 597
|||||
449 ATGTCGTAATTTTCATCAAGGTACTTTAAGGCATATAATCAGACTCACTTGTAGCCTGTGG 508
|||||
598 AACGGGGCTTTTTCATCCCAATTTGACCTACATAGAAATTTGAGCATCATCTCTGAGGACAA 657
|||||
509 AACTGGGGCTTTTTCATCCCAATTTGACCTACATAGAAATTTGAGCATCATCTCTGAGGACAA 568
|||||
658 TATTTTAACTGAGGAGAACTCACAATTTTGAAGAAACGGCCGCTGGGAGAGTCCATATGACCC 717
|||||
569 CATTTTAACTGAGGAGAACTCACAATTTTGAAGAAACGGCCGCTGGGAGAGTCCATATGATCC 628
|||||
718 TAAGCTGCTGACAGCATCCCTTTTAAATAGATGGAGAAATTTATCTCTGGAACCTGCACTGGA 777
|||||
629 CAACTACTGACTGCTCTCTTTTAATAGACGGTGGTGTACTCTGGAACCTGCTGCGGA 688
|||||
778 TTTTATGGGGGAGACCTTTTGTCTATCTTCCGAACTCTTGGGACCAACCAATCAGGAC 837
|||||
689 CTTTATGGGACGGGACTTTTGTCTATCTTCCAGAACTCTGGGGGACCAATCAGGAC 748
|||||
838 AGAGCAGCATGATTTCCAGGTGCTCAATGATCCAAAGTTTCAATAGTCCGACCTCATCTC 897
|||||
749 GGAGCAGCATGATTTCCAGGTGCTCAATGATCCAAAGTTTCAATAGTCCGACCTCATCTC 808
|||||
898 AGAGAGTGACAACTCTGAAAGATGACAAAGTATATCTTTTCTTCCGTGAAATGCAATAGA 957
|||||
809 AGAGAGTGACAACTCTGAAAGATGACAAAGTATATCTTTTCTTCCGTGAAATGCAATAGA 868
|||||
958 TGGAGAACACTCTCTGGAAAAGCTACTCAGCTAGATAGGTGAGATATATCAAGAATGACTT 1017
|||||
869 CGGAGAACACTCTCTGGAAAAGCTACTCAGCTAGATAGGTGAGATATATCAAGAATGACTT 928
|||||
1018 TGGAGGACAGAGAGTCTGGTGAATTAATGAGCAACATTTCTTCAAGCTCGTCTGATTTG 1077
|||||
929 TGGTGGACAGAGAGTCTTGTGAATAAATGAGCAACATTTCTTAAAGACACGCTGATTTG 988
|||||
1078 CTGAGTGCAGGCTCCAAATGGCATTTGACACTCATTTTGTATGAACTGCAAGGATGATTTCT 1137
|||||
989 CTCTGTGCGGCTCCAAATGGCATTTGACACTCATTTTGTATGAAATGCAAGGATGATTTCT 1048
|||||
1138 AATGAACCTTTAAAGATCTTAAAGATCCAGTTGTATATGAGGTGTTTACGACTTCCAGTAA 1197
|||||
1049 AATGAACCTTTAAAGATCTTAAAGATCCAGTTGTATATGAGGTGTTTACAAATCAAGCAA 1108
|||||
1198 CATTTTCAAGGATCAGCGGTGTGTATGATATGAGCATGATGTGAGAAAGGTTTCT 1257
|||||
1109 CATCTTTAAGGATCTGCTGTGTGATGTACAGCATGATGTATGAAAGGTTTCT 1168
|||||
1258 TGGTCCATATCCCAAGGATGAGCAACCACTATCAATGGGTGCTTATCAAGGAAGAGT 1317
|||||
1169 TGGTCCATATGCTACAGAGATGTTCCCACTATCAGTGGGTGCTTACCAAGGAAGAGT 1228
|||||
1318 CCCCTATCCAGGCGCAGGAACTTGTCCAGCAAAACATTTGGTGGTGTGTTGACTTACAAA 1377
|||||
1229 TCCTTATCCAGGCGCAGGAACTTGTCCAGTAAACATTTGGCGGATTTGACTCCACAAA 1288
|||||
1378 GGACCTTCTGATGATGTTTAACTTTTGGAGAGTCCAGGATCCAGGATGATCAATCCAGT 1437
|||||
1289 GGACCTTCTGATGATGTTTAACTTTTGGAGAGTCCAGGATGATCAATCCAGGATGATCAATCCAGT 1348
|||||
1438 GTTTCCTATGAACCAATCGCCCAATAGTATCAAAACGGATGTAATTAATCAATTTACACA 1497
|||||
1349 GTTTCCTATGAATTAATCGCCCGATCATGATCAAAACAGATGTAATTAATCAATTTACACA 1408
|||||
1498 AATTGTGCTAGACCGGAGTGGATGAGAGAGTGGACAGTATGATGTTTATCGGAAC 1557
|||||
1409 AATTGTGCTAGACCGGAGTGGATGAGAGAGTGGACAGTATGATGTTTATCGGAAC 1468
|||||
1558 AGATGTTGGGACCGCTTTTAAAGTATGTTTCAATTTCCCTAAGGAGACTTGGTATGTTTGA 1617
|||||
1469 AGATGTTGGGACCGCTTTTAAAGTATGTTTCAATTTCCCTAAGGAGACTTGGATGTTTGA 1528
|||||

Qy	1618	AGAGGTTCTGCTGGAAGAAATGACAGATTTTTTCGGGAAACCGAGCTGCTATATTTTCAGCAATGGGA	1677
Db	1529	AGAAAGTTTCTTTTCGGAAGAAATGACCGGCTTTCCGGGAAACCAACAACTATTTTCGGCAATGGGA	1588
Qy	1678	GCCTTTCCACTAAGCAGACAAACATATATATTTGGTTCAACCGCTGGGGTTGCCCCAGCTCCC	1737
Db	1589	GCCTTTCTACTTAAACAGCAACAGCTGTACATTTGGCTCAACTTCGGGAGTGGCAGAGTTCC	1648
Qy	1738	TTTACACCGGTGTGATATTTTACGGGAAAGCGTGTGCTGAGTGTGTGCCTCCGCCGAGACCC	1797
Db	1649	TCTACACCGCTGTGACATCTATGGCAAGCCTGTGCAGAAATGCTGCTCCCTCGGAGACC	1708
Qy	1798	TTACTGTGCTTTGGGATGGTTCGTGATGTTCTCGCTATTTTCCACTGCAAGAGACGCAC	1857
Db	1709	TTACTGTGCTTTGGGATGGGTCCCTCATGCTCAACGCTATTTTCTACTGCAAGAGGGGCGAC	1768
Qy	1858	AAGACGACAGATATAAGAAATCGAGACCCACACTGACTCACTGTTCAGACTT---ACACCA	1914
Db	1769	AAGACGACAGATATAAGGAATGGAGACCCACACTGACTCACTGTCTGACTTGGAGGAGCA	1828
Qy	1915	TGATAATCACCATGGCCACACGCCCTGAAGAGAGAAATCATCTATGCTGTGAGAGAAATAGTAG	1974
Db	1829	TGATAATCATCATGGGCCACGCCCTTGAAGAGAGAAATCATCTATGAGTGGAAACAGTAG	1888
Qy	1975	CACATTTTGGAAATGCAAGTCGAAAGTCGAGAGAGCGCTGGTCTATTTGGCAATTCAGAG	2034
Db	1889	TACATTTCTTGGAAATGCAAGTCGAAAGTCGAGAGAGAGATTCAGGCAATTTTCAGAG	1948
Qy	2035	GCGAAATGAAGCGCAAGAGAGATCGAGTGGATGATCATCATCAGGACAGATCA	2094
Db	1949	GAGAAATCGAAGATCGAAAGAGAGAGATCGAAATGGGTGATCATCATCAGGACAGAA	2008
Qy	2095	AGGCCTTCTGCTACGTAGTCTACAACAGAAAGGATTCAGGCAATTTACCTCTGCCATGCGGT	2154
Db	2009	AGGCTCTCTGCTCGTAGCTGCGAAGAGAGGATTCAGGCAATTTACCTGTGTGCTCAGCTGT	2068
Qy	2155	GGAAACATGGGTTTCATAAACTCTTTTAAGGTAAACCTCGAAAGTCATTCAGCACAGAGCA	2214
Db	2069	GGAAACCGGATTCATGCAAACTCTTCTTAAGGTAAACCTCGAAAGTCATTCAGCACAGAA	2128
Qy	2215	TTTGGAGAATCTTCTTATAAAGATGATGAGATGGCTCTTAAGACCAAGAAATGTC	2274
Db	2129	TTTGGAGAATCTTCTTATAAAGATGACGATGGAGATGGCTCTTAAGATAAAGAAATGTC	2188
Qy	2275	CAATPAGCATGACACTTACCGCAGAAGTCTCGTACAGAGACTTCATGCAGCTCATCAACCA	2334
Db	2189	GAGCAGCATGACGCCCGCCAGCCAGAAAGTCTGGTACCGAGACTTCATGCAGCTCATTAACCA	2248
Qy	2335	CCCCAATCTCAACACGATGGATGAGTTCCTGTGAACAAAGTTTGGAAAGGACCGGAAACCA	2394
Db	2249	CCCCAATCTGAAACACGATGGATGAGTTCCTGTGAACAAAGTGTGGAAAGGACCGGAAACCA	2308
Qy	2395	ACGTGCGCAAGGCCAGGACATACCCAGGGGACAGTAAACAAATGGGAAGCACTTACAGA	2454
Db	2309	ACGCCGCAAGGCCGGGGCACTCTCAAGGGAGCAGCAACAAAGTGGGAAGCACATGCAAGA	2368
Qy	2455	AAATAAGAAAGGTAGAAACAGGAGGCCACAGAAATTTGAGAGGCGCACCCAGGAGTGTCTG	2514
Db	2369	GAGCAAGAAAGGTAGAAACAGGAGGCCACAGAGTTTGGAGGCGCACCCAGGAGTGTCTG	2428
Qy	2515	AGCTGCTATTACCTCTAGAAACCTCAAAACAAAGTAGAACTTTGCCCTAGACAAATAACTGGAAA	2574
Db	2429	AGCTGCGCCACCTCCCAAAACCTCAAAACAAAGTACAACTTGTCTTAGATAATACTGGAAA	2488
Qy	2575	AACAAATGCAATATACATGAATCTTTTTCATGGCAATTTATGGATGTTTACAATGGTGGG	2634
Db	2489	---AAATGCAATACACATGAAC--ATTTCATGGCAATTTATGGATGTTTACAATGATGGG	2543
Qy	2635	AAATTCAGCTGAGTTCCACCAATTTATAAATTAATCCATGAGTAACTTTCCCTAATAGGCT	2694
Db	2544	AAGTTCAACCGGGTTCCACCAAT-----TTAAGTCGTGAGGAACTTTTCCAGCAGGCT	2597
Qy	2695	TTTTTTTCCCTA	2704

Df 2598 TTCTTCAGTA 2607

RESULT 10
RNSIIICNI RNSIIICNI 2319 bp mRNA linear ROD 29-JAN-1996
LOCUS R.norvegicus mRNA for semaphorin III/collapsin-1.
DEFINITION X95286
ACCESSION X95286.1 GI:11717618
VERSION collapsin-1; semaphorin III.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
1 Giger,R.I.
AUTHORS Unpublished
JOURNAL 2 (bases 1 to 2319)
REFERENCE Giger,R.I.
AUTHORS Direct Submission
TITLE Submitted (22-JAN-1996) R.I. Giger, Netherlands Inst. for Brain
JOURNAL Research, Meibergdreef 39, 1105 AZ, Amsterdam, NETHERLANDS

FEATURES
source
1..2319
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Wistar"
/db_xref="taxon:10116"
/tissue_type="brain"
/dev_stage="adult"
1..2319
/codon_start=1
/product="semaphorin III/collapsin-1"
/protein_id="CAA64607.1"
/db_xref="GI:1171619"
/db_xref="SPTREMBL:Q63548"
/translations="MGVFTGIACLFWGILLTARANYANGQNVPRLKLSYKMELENS
VITFNLGNSSSVHTFLDEERSLVGAKHIFSNLVNIDKFQKIWMPSVTREDE
CKWAGXDLKECANFIKVLKYNQTHLYACGTGAFHPDICTIYEVGHPEDNIRFKLODS
HFNFGKCDYSPLDKLLTASLDIGELYSGTDAFMGRDFAIPTLGHHPIRTREQHDS
RWLNDRPFTSAHLIPSDNPDDKVYFFPRENAIDGEHSKATHARIQVICKNDFGSNI
RSLVNRKWTTFPKARLKICSVPNGNDITHDELQDVLFMLMSKDPKNPIVIYGVFTTSNI
FKGSAVMTSMDSVRVFELGPAYHRDGPNYQWPYQGRVYPFRPGTSPKSTFCGPDSI
KDI.PDDVLTIFARSHPAVMYVPFINNRPIMIKTDVNYQFTOIWDVDAEDGYDVMF
IGIDVGTVLKVSVPKETWHDLLEUVLEEMTVPRETTISAMELSTKQQOYLIGSTAG
VAQLPHRCDDIYGKCAECCCLARDPYCAMWDCSCSRFFPTAKRRTRRDRIIRNGDPLTH
CSDLQHNNHHGHSLERI.YGVENSTFLECPKSORALVYWQFORNEERDKBEIRI
GDHI IRTQEOLSKSYOKWYNLCHEAVEHGMPQLTKVKTELEVIDTHELEELLHKDD
DGGSTKTCKWSSMTPSKKVYHDVFMQLINHNLNTMDFECEQVWKRRDRKQRQRPRPH
SQSSNNKWKMQESKGRNRRTHEFERAPRV"

BASE COUNT 704 a 491 c 548 g 576 t
ORIGIN

Query Match 70.5%; Score 1909.4; DB 10; Length 2319;
Best Local Similarity 89.3%; Pred. No. 0;
Matches 2070; Conservative 0; Mismatches 246; Indels 3; Gaps 1;

Qy 200 ATGGCGTGTTAACTAGGATGTCTGCTTTTTCTGGGAGTATTACTTACAGAAGCA 259
| | | | |
Db 1 ATGGCGTGTTCACTGGGATTCCTGCTCTTTCTGGGTATATTACTTACAGAAGCA 60
| | | | |

Qy 260 AACTATCAGAAATGGGAAGACAATGTGCCAAGCTGAATATTCTTACAGAAGATGTTG 319
| | | | |

Db 61 AACATATGCAACAGGAAGAACAAATGTGCCAAGACTGAAATATTCTGACAGAAGATGTTG 120
| | | | |

Qy 320 GAATCCAAACAATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCATCTTC 379
| | | | |

Db 121 GAATCCAAACAATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCATCTTT 180
| | | | |

Qy 380 CTTTGTGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAGGATCACATATTTTCATTC 439
| | | | |

181	CTTCTGGATGAGGAACGGAGTAGACTGTATGTTGGAGCAAAAGGATCATATATATTTTCATTTC	240
440	GACCTGGTTAATATCAAGAGATTTTCAAAAGATTTGTGGCCAGTATCTTACACCAAGAGA	499
241	AACTTGGTGAATATTAAGATTTTCAAAAGATTTGTGGCCAGTATCTTATACCAAGAGA	300
500	GATGAATGCAAGTGGCTGGAAAAAGACATCTCTGAAAGAAATGTGTCTAATTTTCATCAAGGTA	559
301	GATGAATGCAAGTGGCTGGAAAAAGATATCTGTAAAGAAATGTGTCAAAATTTTCATCAAGGTC	360
560	CTTAAGGCATATATCAGATCTGATCGCCTGTGGACGGGGGCTTTTTCATCCCAATTT	619
361	CTCAAGGCTTATATCAGACTCATCTGTGACCTGTGTGAAACGGGGGCTTTTCCATCCCAATTT	420
620	TGCACCTACATTTGGAATTTGGACATCATCTCTGAGGACAAATATTTTAAAGCTGGGAGAACTCA	679
421	TGCACCTATATTGAAGTTTGGACATCTCTGAGGACAAATTTTAACTTGCAGGACTCA	480
680	CATTTTGAAAAACGGCCGTGGGAAGAGTCCATATGACCCCTAAGCTGTCTGACAGCATCCCTTT	739
481	CATTTTGAAAAACGGCTGTGGGAAGAGTCTTATGACCCCAAACTTCTGTGACTGCTCTCTTT	540
740	TTAATAGATGGAGAAATTAATCTCTGGAACTGACAGTCAATTTATTGGGGCGAGACTTTGCT	799
541	CTAATAGATGGTGAATTTATCTCTGGAACCTGCTGCGAATTTTCATGGGACAGAGACTTTTGT	600
800	ATCTTCCGAACCTCTTTGGGCACCAACCCCAATCAGGACAGAGCAGCATGATTCACAGTGG	859
601	ATCTTTCAGAACATTTGGGGCACCATCTCCATCAGGACCGAGCAGCAGACTCCCGGTGG	660
860	CTCAATGATCCAAAGTTCAATTTAGTGGCCCACTCATCTCAGAGAGTGACAATCTCTGAAGAT	919
661	CTCAATGATCTAGATTTTCATCAGTGGCCCATCTCATCCAGAGAGTGATAACCCCTGAAGAT	720
920	GACAAAGTATACTTTTCTCCGTGAAAAATGCAATAGATGAGAGAACTCTGTGAAAAGCT	979
721	GACAAAGTATATTTTTTTCTTCCGAGAAAAATGCAATAGATGAGAGAACTTCTGTGAAAAGCC	780
980	ACTCACGCTAGAATPAGTTCAGATATCAAGAAATGACITTTGAGGGGACACAGAGTCTGGTG	1039
781	ACTCATGCTAGAATPAGTTCAGATATCAAGAAATGACITTTGAGGGGACACAGAGTCTGGTG	840
1040	AATAAATGGACAACTTCTCCTCAAAAGCTCGTCTGATTTTGTCTCAGTGCCAGGTCCTCAAAATGGC	1099
841	AATAAATGGACAACTTCTCCTCAAAAGCAGCCTGATTTGCTCTGTGCCAGGTCCCAATGGC	900
1100	ATTGACACTCATTTTGTGATGAATCTGAGAGTGTATTCCTTAATGAACTTTTAAAGATCTCTAAA	1159
901	ATTGACACACATTTTGTGATGAATTTACAGGATGTATTTCTTAATGAACTCTTAAAGATCTCTAAA	960
1160	AATCCAGTGTATATGAGTGTTTTACCACTTCCAGTAACTATTTTCRAAGGATCAGCGGTG	1219
961	AATCCAAATGTCTATGAGTGTTCACAACTATCAAGCAAACTATTTTAAAGGCTCTGTGTG	1020
1220	TGTATGTATAGCATGATGTGTGAGAAGGCTGTCTCTGGTCCATATGCCACAGGAT	1279
1021	TGTATGTACAGTATGATGTATGAGAAGGCTGTCTCTGGTCCATATGCTCACAGAT	1080
1280	GGACCCCAACTATCAATGGGTGCCCTTATCAAGGAAGAGTCCCTTATCCAGCCAGGAACT	1339
1081	GGTCCCAACTATCAGTGGGTACCTTACCAAGGAAGAGTCCCTTATCCAGCCAGGAACT	1140
1340	TGTCCCAAGCAAAACATTTTGGTGGTTTGTACTCTCAAGAGCACTTCTGTGATGTATATA	1399
1141	TGTCCCAAGTAAACATTTTGGTGGATTTGACTCTCAAAAGAACCTTCTGTGATGTATATA	1200
1400	ACCTTTTCCAAAGAAGTCAATCCAGCCATGTCAAACTCCAGTGTCTTCTATGAACAATTCGCCCA	1459
1201	ACCTTTTGCAGAGAGCCATCCAGCCATGTCAAACTCCAGTGTCTTCTATTAATATTCGCCCA	1260
1460	ATAGTGATCAAAACGGAGTGAATTTATCAATTTTACAAATTCGATGAGACCGAGTGGAT	1519
1261	ATAATGATCAAAACAGAGTGAATTTATCAATTTTACAAATTTGTGTAGACCGGGTGGAT	1320

QY	1520	GCAGAGATGGA	CAGTATGATGATGTTATTCGGAA	CAGATGTTGGGACCGGTTCTTAA	1579
DB	1321	GCAGAGATGGC	CAGTATGATGTTATGTTTCATCGAA	CGGATGTTGGAACCTGTCCTTAA	1380
QY	1580	GTAGTTTCAAT	TCTCTAAGGAGACTTGATATGATTTAGAA	GAGGTTCTGCTGGAGAAATG	1639
DB	1381	GTGGTTTCAGT	CCCCAAGGAGACTTGGCATGACTTAGA	AGAAGTTCTTCTGGNAGAAATG	1440
QY	1640	ACAGTTTTTTC	GGGAACCGACTGCTATTTTCAGCAAT	GGAGCTTTCCAATAAGCAGCAACAA	1699
DB	1441	ACAGTCTTCT	CGGGAACCAACAACTATTTTCAGCAAT	GGAGCTTTCTACTAAACAGCAACAA	1500
QY	1700	CTATATATTGG	TTCACGGCTGGGGTTGCCAGCTCCCTTTAC	ACCGGTGTGATATTTTAC	1759
DB	1501	CTGTACATTT	GGGCTCAACTGCTGGAGTAGCACAGCTTCCT	CTACACCGTTGTGACATTTAT	1560
QY	1760	GGGAAACGGT	GTGCTGAGTGTGCCCTCGCCGAGAGCCCTTACT	TGTGCTTTGGGATGTTCT	1819
DB	1561	GGCAAGCCTG	TGGGAACTGCTGCTTGTCTGGGACCTTATTTG	TGCTTGGGATGGGTCA	1620
QY	1820	GCATGTTTCT	CGCTATTTTCCCATGCAAAAGAGACGCA	CAAGACGACAAGATATAAGAAAT	1879
DB	1621	TCATGCTCT	ATTTTCTCTACTGCAAAAGAGCGCACAA	GACGACAGGATATAAGGAAAT	1680
QY	1980	GGAGACCACT	GACTACTCTGTTTCAGACTT---ACACCAT	GATTAATCACCATGCCACAGC	1936
DB	1681	GGAGACCACT	GACTACTCTGCTCCGACTTCGACATCAT	GATTAATCATCATGGCACAGT	1740
QY	1937	CCTGAAGAGA	GAATCATCTATGTGTAGAGAAATAGTAG	CACATTTTTTGGAAATGCAGTCG	1996
DB	1741	CTTGAAGAGA	GAATCATCTATGAGTAGAAATAGTAG	CACATTTCTTGGAAATGCAGTCG	1800
QY	1997	AAGTCGCGAG	AGCGTGTCATTTGGCAATTTCCAGAGG	CGAAATGAAGAGCGAAAGAA	2056
DB	1801	AAGTCACAGA	GAGCCTTTGGTATATTGGCAATTTT	CAGAGCGCAATGAAGATCGAAAGGAG	1860
QY	2057	GAGATCAGAG	TGGATCATATCATCAGCAGACAGATCA	AGGGCTTCTGCTACGTAGTCTTA	2116
DB	1861	GAGATCAGAG	TGGGTGATCATATCATCAGCAGACAA	CAGGGGCTCTGCTTCGTAGTCTTG	1920
QY	2117	CAACAGAAGG	ATTTTCAGGCAATTTACCTCTGCGAT	GCGGTGGAAACATGGGTTTCATACAAACT	2176
DB	1921	CAGAAAGAAG	ATTTTCAGGCAATTTACCTGTCTACG	CGGTGGAGCAGGTTTTCATGCAAACT	1980
QY	2177	CTTCTTAAAG	GTAAACCTCTGGAAGTCATTTGAC	CAGAGCATTTTGGAGAACTTCTTCTATAAA	2236
DB	1981	CTTCTTAAAG	GTGACCTCTGGAAGTCATTCGATAC	CAGAACTTTTGGAGGAACTTCTTCTATAAA	2040
QY	2237	GATGATGATG	GGAGTGGCTCTTAGACCAAAAGAAAT	TGTCCAATAGCATGACACTAGCCAG	2296
DB	2041	GATGACGATG	GGAGATGGCTCTTAGACCAAAAGAAAT	TGTCAGTAGCATGACGCCAGCCAG	2100
QY	2297	AAAGTCTGGT	AACAGACTTTCATGTCAGCTCATCAAC	CCCCCAATCTCAACACGATGGAT	2356
DB	2101	AAAGTCTGGT	AACAGACTTTCATGTCAGCTCATCAAC	CCCCCAACCTGGAACAATGGAT	2160
QY	2357	GAGTTCTGTG	AAACAGATTTTGGAAACGGGACCGGAA	AAACAACTGTCGGCAAGGCCAGGACAT	2416
DB	2161	GAGTTCTGTG	AAACAGATTTTGGAAACGGGACCGGAA	AAACAAACGCGGCAACAAAGGCCGGGGCAC	2220
QY	2417	ACCCCGGAA	CAGTAAACAAATGGAAGCACTTAC	AGAAAAATAAGAAAGGTAGAACAGG	2476
DB	2221	TCTCAAGG	GAGCAGCAACAGTGAAGCACTG	CAGAGAGCAAGNAAGGTAGAACAGG	2280
QY	2477	AGGACCCAG	AAATTTTGAGGGGCCACCCAGGAGT	GTCTGA	2515
DB	2281	AGGACCCAG	AAATTTTGAGGGGCCACCCAGGAGT	GTCTGA	2319

RESULT 11
MUSCOSE
LOCUS

MUSCOSE	2319 bp	mRNA	linear	ROD 13-FEB-2002
---------	---------	------	--------	-----------------


```

DEFINITION Mus musculus collapsin/semaphorin III mRNA, complete cds.
ACCESSION L41541
VERSION L41541.1
KEYWORDS collapsin/semaphorinIII.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 2319)
AUTHORS Kimura,T. and Fishman,M.C.
TITLE cDNA sequence of mouse Collapsin/SemaphorinIII
JOURNAL Unpublished
FEATURES
    Location/Qualifiers
        1..2319
            /organism="Mus musculus"
            /mol_type="mRNA"
            /db_xref="taxon:10090"
            /dev_stage="adult"
        1..2319
            /note="putative"
            /codon_start=1
            /product="collapsin/semaphorin III"
            /protein_id="AA17761.1"
            /db_xref="GI:18654370"
            /translation="MGWFTGIACLFMGVLLITARANYANGKNVPLKLSYKEMLESN
            VITFGLANSSTYHTFLDLDEBSRLVYGAKHIFSNLVNINIKDFQIVPVSYYRDE
            KWAGKDLKECANFTKVLAEYNQTHLYACGTGAGHPICITYIEVGHPEPDIKFLQDS
            HPENGRKSPYADPKLLTASLLDGLSYGTAADFMGRDPAIFRPLTGHHPHPIRTEOHS
            RLWNPRTISAHLLIPESNPEDDKVYFFRENAIDGHSKATARIQICIKNDGCGH
            RSLNKNWTFKLARLICSVPNGIDITHDELQVFLMNSDKPNPIVIGVFTTSNI
            FKGSACVMTSDRRVFLGPAHYHRDGNYPQVYGRVPYPRPTGCTSKTGGFDDST
            KDLPDVITFARSHPAMTNPVPIINRIMIKTDVNYQTOIVVDVDAEDGQYDMF
            IGTDVTLKVVSVKPEYTHLDEVLLEWTFREPTTISAMELSTKQQLYIGSTAG
            VAQLPLRCDIYGAACECLLDPCYCAWDSSCSRYPTAKRRTRRDIRNGDPLTH
            CSDLQHNHNGHPSLEERIIYGVENSSTFLECSPEKORALVYVQFQRENDRKEIRM
            GDHLIRTEQGLLRSLQKDSGNLYLCHAVEHGMOTLLKLVILEVIDTHELLEHLLKOD
            DGGSKIEMSSMTSPQKVTYRDFMQLNHNPLNTMDEFCEQVWKRORRORRPHG
            SGGSNKKWQHMQESKKGNRRTHFERAPRSV"
BASE COUNT 703 a 513 c 549 g 554 t
ORIGIN
    Query Match 69.5%; Score 1883.8; DB 10; Length 2319;
    Best Local Similarity 88.6%; Pred. No. 0;
    Matches 2054; Conservative 0; Mismatches 262; Indels 3; Gaps 1;
QY 200 ATGGGCTGTTAACTAGGATTGTCTCTTTCTGGGGAGTATTACTTACAGCAAGAGCA 259
DB 1 ATGGGCTGTTCACTGGGATTCCTCTTTCTGGGGTGTATTACTTACGCCAGAGCA 60
QY 260 AACTATCAGAAATGGGAAGAACAAATGTGCCAAGCTGAAATATCTTACAAAGAAATGTTG 319
DB 61 AACTATGCAACGGGAAGAACAAATGTGCCAAGCTGAAATATCTTACAAAGAAATGTTG 120
QY 320 GAATCCAAAGTGTATCACTTTCAATGGCTTGGCCAAAGCTCAGATTATCATCCTTC 379
DB 121 GAATCCAAAGTGTATCACTTTCAATGGCTTGGCCAAAGCTCAGATTATCATCCTTC 180
QY 380 CTTTGGATGAGGAAGGAGTGGCTGTATGTTGGAGCAAGGATCACATATTTTCATTC 439
DB 181 CTTTGGATGAGGAAGGAGTGGCTGTATGTTGGAGCAAGGATCACATATTTTCATTC 240
QY 440 GACCTGGTTTAAATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACACCAAGA 499
DB 241 AACTTGGTGAAACATTAAAGATTTCAAAAGATTTGTGGCCAGTATCTTACACCAAGA 300
QY 500 GATGAATGAAGTGGCTGGAAAGACATCTTGAAGAAGATGCTTAATTTCAATCAAGTTA 559
DB 301 GATGAATGAAGTGGCTGGAAAGATCTTCTGAAAGATGTGCCAATTTTCATCAAGTTC 360
QY 560 CTTAAGGCATATAATCAGACTCCTGTGTACGCTGTGGAAACGGGGCTTTTCATCAAT 619
DB 361 CTGAGGCTTATAATCAGACTCCTGTGTATGCTGTGGAACTGGGGCTTTTCATCAATC 420

```

620 TGCACCTAATTGAAATTTGGACATCATCTCTGAGGACAAATATTTTTTAAGCTGAGAACTCA 679

421 TGCACCTAATTGAAATTTGGACATCATCTCTGAGGACAAATATTTTTTAAGCTGAGAACTCA 480

680 CATTTTGAACGGCGCTGGGAAGAGTCCATATGACCTTAAGCTGCTGACACATCCCTT 739

481 CATTTTGAACGGCGCTGGGAAGAGCCTTATGATCCAAACTACTGACTCCCTCTCTT 540

740 TTAATAGATGAGAAATATTACTCTCTGAAACTGACGCTGATTTTTTATGGGGGAGACTTTGCT 799

541 CTAATAGACGGTGTGCTCTCTGAACTGCTGGGACTTCTATGGGACGGGACTTCGCT 600

800 ATCTTCCGAACTCTTGGGACCAACCAATCAGACAGACAGACAGATGATTCAGGTGG 859

601 ATCTTCCGAACTCTTGGGACCAATCAGGACAGACAGACAGATGATTCAGGTGG 660

860 CTCAATGATCAAAAGTCTCATTAAGTCCCACTCATCTCAGAGAGTGAACAATCTCAAGAT 919

661 CTCAATGATCCTAGATTCATCAGTCCCATCTCATCCAGAGAGTGAACAATCTCAAGAT 720

920 GACAAAGTATATCTTTCTTCCGTGAAATGCAATAGATGGAGAACTCTGGAAGGCT 979

721 GACAAAGTATATTTTCTTCCGAGAAATGCAATAGACGGGAGAACATCTGGAAGGCT 780

980 ACTACGCTAGATAGTGTAGATATGCAAGAAATGATCTTTGGAGGACAGAAATCTGCTG 1039

781 ACTCATGTAGATAGTGTAGATATGCAAGAAATGATCTTTGGTGGACAGAAATCTGCTG 840

1040 AATAAATGGAACAATCTCTCAAGCTCGTCTGATTTGCTCAGTCCGAGGTCGAATGGC 1099

841 AATAAATGGAACAATCTCTCAAGCTCGTCTGATTTGCTCAGTCCGAGGTCGAATGGC 900

1100 ATTGACACTCATTTTGTGAACTGAGGATGTATTCTTAAGAACTTTAAAGATCTCTAAA 1159

901 ATTGACACTCATTTTGTGAACTGAGGATGTATTCTTAAGAACTTTAAAGATCTCTAAA 960

1160 AATCAGTTGTATATGAGTGTATGAGTCTTCAAGTCTCAGTAAATTTTCAAGGATCAGCGTG 1219

961 AATCAGTGTATGAGTGTATGAGTCTTCAAGTCTCAGTAAATTTTAAAGGATCTGCTGTG 1020

1220 TGTATGATAGTATGATGATGATGAGAGGTTGTTCTTGTCCATATGCCCCACAGGAT 1279

1021 TGTATGATGATGATGATGATGAGAGGTTGTTCTTGTCCATATGCTCAGAGAT 1080

1280 GGACCCAACTATCAATGGGCTCTTATCAAGAAAGAGTCCCTTATCCAGCGGACAGAACT 1339

1081 GGTCCCACTATCAGTGGTCTTACCAAGAGAGTCCCTTATCCAGCGGACAGAACT 1140

1340 TGTCCCAAGAAACATTTGGTGGTTCATCAAAAGGACCTTCTGATGATGTTATA 1399

1141 TGTCCCAAGAAACATTTGGTGGTTCATCAAAAGGACCTTCTGATGATGTTATA 1200

1400 ACCTTTGCAAGAGTCAATCCAGCCATGATCAATCCAGTGTTCCTTATGAACATCGCCA 1459

1201 ACTTTGCAAGAGTCAATCCAGCCATGATCAATCCAGTGTTCCTTATGAATCGCCG 1260

1460 ATAGTGTCAAAACGGATGTAATTTATCAATTTTACAAATTTGCTAGACCGAGTGGAT 1519

1261 ATCATGATCAAAACAGATGTAATTTATCAGTTTCACAAAATTTGTTAGACCGAGTGGAT 1320

1520 GCAGAAGATGACAGTATGATGTTATGTTTATCGGAAACAGATGTTGGACCGTCTTAAA 1579

1321 GCAGAAGATGACAGTATGATGTTATGTTTATCGGAAACAGATGTTGGACCGTCTTAAA 1380

1580 GTAGTTTCAATTTCTAAGGACATGTTGATGATTTAGAGAGGTTCTCTCTGGAAGAAATG 1639

1381 GTGGTTTCAGTCCCCAAGGAGACTTGGCATGACCTTAGAAGAGTCTTCTCTGGAAGAAATG 1440

1640 ACAGTTTTTTCGGGAACCGACTGCTATTTTCACAATTTGAGGCTTCCACTAAGCAGCAAA 1699

1441 ACCGTTCTCCGGGAACCAACAACTATTTTCGGCAATGAGGCTTTCTACTAAACAGCAACAG 1500

1700 CTATATATTGTTCAACGGGCTGGGGTGGCCAGCTCCCTTTACACCGGCTGTGATATTTAC 1759

```

1501 CTGTACATTGGCTCAACTGGCGAGTGGCACAGCTTCTCTACACCGCTGTGACATCTAT 1560
1760 GGGAAAGCGTGTGCTGAGTGTTCCTCGCCGAGACCCCTTACTGTGCTTGGATGTTCT 1819
1561 GCGAAAGCGTGTGACAAATGCTGCTCGCTCGCGGACCCCTTACTGTGCTTGGATGTTCT 1620
1820 GCATGTTCTCGCTTATTTTCCACTGCAAGAGAGACGACAGACGACGACGACGACGACGAC 1879
1621 TCATGCTACGCTTATTTTCTACTGCAAGAGCGGCAAGACGACGACGACGACGACGACGAC 1680
1880 GGAGACCCACTGACTGCTTCTGACTT---ACACCATGATGATGATGATGATGATGATGATGAT 1936
1681 GGAGACCCACTGACTGCTTCTGACTTGCAGCACCACGATGATGATGATGATGATGATGATGAT 1740
1937 CCTGAAGAGAGATCATCTATGCTGTAGAGATGATGATGATGATGATGATGATGATGATGAT 1996
1741 CTTGAAGAGAGATCATCTATGAGTGGGAAACAGTAGTACATCTTGGAGTGGAGTCCG 1800
1997 AAGTCGACAGAGCGCTGCTTATTTGGCAATTCAGAGCGGAAATGAGAGCGGAAAGAA 2056
1801 AAGTCACAGAGCGCTTGGTATTTGGCAATTTTCAAGAGGAGAAATGAGATGAAAGAG 1860
2057 GAGATCAGAGTGGATGATCATATCATCAGACAGATCAAGGCGCTTCTGCTAGCTGATCTA 2116
1861 GAGATCAGAGTGGTATCATATCATCAGACAGAGAGGCGCTTCTGCTAGGCTG 1920
2117 CAACAGAGGATCAGGCAATACCTTCTGCTGCGTGGAGACATGCGTTTCAACAAT 2176
1921 CAGAAGAGGATTCAGGCAATACCTTCTGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1980
2177 CTTCTTAAGTAACTGAGGAGTCAATGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2236
1981 CTTCTTAAGTAACTGAGGAGTCAATGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2040
2237 GATGATGATGAGGAGGCTTAAAGACCAAGAAATGCTCAATGAGGAGGAGGAGGAGGAGGAGGAG 2296
2041 GATGACGATGGAGTGGCTTAAAGATAAAGAAATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 2100
2297 AAGTCTGTGACAGAGCTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2356
2101 AAGTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2160
2357 GAGTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2416
2161 GAGTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220
2417 ACCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2476
2221 TCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2280
2477 AGGACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2515
2281 AGGACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2319

```

RESULT 12

GGU02528 3263 bp mRNA linear VRT 01-NOV-1993

Gallus gallus collapsin mRNA, complete cds.

DEFINITION
U02528
VERSION
U02528.1 GI:410078

Gallus gallus (chicken)

ORGANISM

Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

Phasianinae; Gallus.

1 (bases 1 to 3263)

Luo, Y., Raible, D. and Raper, J. A.

Collapsin: a protein in brain that induces the collapse and

paralysis of neuronal growth cones

Cell 75 (2), 217-227 (1993)

JOURNAL

MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

94006554

8402908

2 (bases 1 to 3263)

Luo, Y.

Direct Submission

Submitted (13-OCT-1993) Yuling Luo, Department of Neuroscience,
University of Pennsylvania, School of Medicine, Philadelphia, PA
19104 USA

FEATURES

Location/Qualifiers

source

1..3263
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/sex="female"
/tissue_type="brain"
/clone_lib="chick brain cdna library (Clontech Cat.
#CL1016A)"
/dev_stage="adult"
168..2486
/codon_start=1
/product="collapsin"
/protein_id="AAC59638.1"
/db_xref="GI:410078"

CDS

1007 a 681 c 713 g 862 t
BASE COUNT
ORIGIN

Query Match 64.4%; Score 1744; DB 5; Length 3263;
Best Local Similarity 81.4%; Pred. No. 0;
Matches 2047; Conservative 0; Mismatches 465; Indels 4; Gaps 2;
QY 189 GCGTCGACGATGGCTGGTTAACTAGGATGCTGCTTCTCTGGGAGTATTACTTGA 248
DB 157 GCTGCTGACGATGGCTGGCTCAGGGGATGGCTTCTTCTTCTGGAGATTACTAG 216
QY 249 CAGCAAGAGCAAACTATCAGAAATGGAAAGAAACAATGTGCCAAGGCTGAAATATCTCTACA 308
DB 217 CAGGAGAGTGAACCTGCCAGCATGTGAAGATAAGTGTCCAAGACTCAAAATATCTCTATA 276
QY 309 AAGAAATGTTGGAATCCAACAATGTGATCACTTTCAATGGCTGGCCCAACAGCTCAGTT 368
DB 277 AAGAAATGTTGGAATCCAAATATAGTCAATTTCAATGGACTAGTCAATAGTTCAGTT 336
QY 369 ATCATACCTTCTTCTGGATGAGGACGAGTGGCTGATGTTGGAGCAAGGATCTACA 428
DB 337 ACCATCTTCTCTTGGATGAAGCAAGCTGGCTGTATGTTGGAGCAAGAAAGACCACA 396
QY 429 TATTTTCAATCGACTGCTTAAATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTT 488
DB 397 TATTTTCTTCAACCTGTTAAACATCAAGGATATCAAAAGATTTGTGGCCCTGTATCTC 456
QY 489 ACACGAGAGAGATGAATGCAAGTGGCTGGAAAGACATCTGAAAGATGTGCTAATT 548
DB 457 ATTCCGAGGAGTGAAGTGAAGTGGCTGGAAAGATATTTCTGAGAAATGTGCTAATT 516
QY 549 TCATCAAGTACTTAAGGCATATATACAGCTCACTTGTACGCTGTGAGAGGGGGCTT 608
DB 517 TCATCAAGTGTCTTAAGACTTATACCAACACACACTTGTACGCTTGTGAGAGGGGGCTT 576
QY 609 TTCATCCAAATTTGCACCTTACATTTGAAATTTGACATCATCTCTGAGGACAAATATTTTAAAGC 668

Db 577 TTCAATCCCATGTCACCTTATATGAGTTGGAAGCCATCTCAGAGCAACATTTTATGGA 636
QY TGGAGAACTCACATTTTGAACCGCGTGGGAAGAGTCCATATGACCCCTTAAGCTGCTGA 728
Db TGGAGAACTCACATTTTGAACCGCGTGGGAAGAGTCCATATGACCCCTTAAGCTGCTGA 696
QY CAGCATCCCTTTTATAGATGAGAAATATATCTCTGGAACTCGAGCTGATTTTATGGGCG 788
Db CAGCTTCTCTTTTAGTAGATGAGAAATGTTACTCTGGCACAGCAGACATTTTATGGGCA 756
QY GAGACTTTGCTATCTTCCGAACTCTTTGGGCACACACCCCAATCAGGACAGAGCAGCATG 848
Db GAGACTTTGCTATCTTCCGAACTCTTTGGGCACACACCCCAATCAGGACAGAGCAGCATG 816
QY ATTCAGAGTGGCTCAATGATCAAAAGTTCAATGATGCCCCACCTCATCTCAGAGAGTGACA 908
Db ACTCCCGGTGGCTAAACGATCCAGATTTTATGAGTCTCACCTGATACCTGAAAGTGATA 876
QY ATCTGGAAGATGACAAAGATATATTTTCTCCGTTGAAATGCAATGATGAGAGAACT 968
Db ATCCGAGAGATGACAAATCTATTTCTCTTTCTGTAATGCGATTTGATGAGAGCACA 936
QY CTGGAAGAGCTACTCACGCTAGAAATAGTGCAGATATGCAAGATGACCTTTGGAGGGACA 1028
Db CTGGAAGAGCTACTCACGCTAGAAATAGTGCAGATATGCAAGATGACCTTTGGAGGGACA 996
QY GAGGCTCTGCTGTAATGAAATGGAACAATCTTCCCAAGCTCGTCTGATTTGCTCAGTGCCAG 1088
Db GAGGCTCTGCTGTAATGAAATGGAACAATCTTCCCAAGCTCGTCTGATTTGCTCAGTGCCAG 1056
QY GTCCAAATGGGCTGACACTCATTTTGTATGAACTGCAAGATGATTCCTAATGAACTTTA 1148
Db GACCAATGGTATGACACACACTTTGATGAAATGACAGACGCTGTTCTAATGAACTCTA 1116
QY AAGATCTTAAATCAATCGATGATGAGTGTATGAGTGTATGAGTGTATGAGTGTATGAGT 1208
Db AGGACCTTAAATCAATCGATGATGAGTGTATGAGTGTATGAGTGTATGAGTGTATGAGT 1176
QY GATCAGCGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1268
Db GATCAGCAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1236
QY CCCACAGGATGAGCCCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1328
Db CCCACAGGATGAGCCCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 1296
QY GGCAGGAACTGTCCAGCAAAACATTTGGTGGTTTGTGATCTCAAGAGACCTTCTG 1388
Db GGCAGGAACTGTCCAGCAAAACATTTGGTGGTTTGTGATCTCAAGAGACCTTCTG 1356
QY ATGATGTTTAACTTTGCAAGAGTCAATCCAGCCATGTACAACTCCAGTGTTCCTATGA 1448
Db ATGATGTTTAACTTTGCAAGAGTCAATCCAGCCATGTACAACTCCAGTGTTCCTATGA 1416
QY ACAATGCCCCAATGATGATCAAAACGAGTGAATTAATCAATTTTACAAATTTGCTGAG 1508
Db ACAGCGGTCCCAATCATGATCAAAACGAGTGAATTAATCAATTTTACAAATTTGCTGAG 1476
QY ACCGAGTGGATGCAAGATGACAGATGATGATGATGATGATGATGATGATGATGATGATG 1568
Db ATCGAGTGGATGCAAGATGCAAGATGATGATGATGATGATGATGATGATGATGATGATG 1536
QY CCGTCTTTAAAGTGTGTTCAATTTCTAAGGAGACTTGGTATGATTTTGAAGAGTTCG 1628
Db CAGTCTTTAAAGTGTGTTCCATCTCTAAGGAGACTTGGATGATTTGAAGAGTTCG 1596
QY TGGAGAAATGACAGTTTTTCCGGAACCGACTGCTATTTTCAGCAATGAGCTTTCCACTA 1688
Db TGGAGAAATGACAGTTTTTCCGGAACCGACTGCTATTTTCAGCAATGAGATTTCCACA 1656
QY AGCAGCAACACTATATATGTTTCAAGGCTGGGTTGCCAGCTCCCTTTACACCGGT 1748

Db 1657 AACGCAACAGCTCTACATTTGGCTCAGCCACGGAGTTTTCAGAGTTCCTTTACACCGCT 1716
QY GTGATATTTACGGGAAGCGTGTGCTGAGTGTGCTCGCCCGAGACCCCTTACTGTGCTT 1808
Db GTGATGTCTATGGAAGAACATGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1776
QY GGGATGTTCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1868
Db GGGATGTTCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1836
QY ATATAAGAAATGAGAGACCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1925
Db ACATCCGAATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1896
QY ATGGCCACAGCCCTGAAGAGAGATCATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1985
Db GTGGTCAAGCCCTGAGGAGAAAGATTTATGAGTAGAGATAGAGAGACCTTTTCTG 1956
QY AATGAGTCCGAAGTCCGAGAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2045
Db AGTGAGTCCGAAGATCTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2016
QY AGCGAAAGAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTGCT 2105
Db ACCAAGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2076
QY TACGTAGTCTACAACAGAGGATTCAGGCAATTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2165
Db TACGAAGTCTTCAACGAGAGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2136
QY TCATCAAACTCTTCTTAAAGTAACTTGAAGTCAATTTGACACAGAGCAATTTGGAAGAAC 2225
Db TCTTCTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2196
QY TCTTCTAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2285
Db TGCTCATAG 2256
QY CACTAGCAG 2345
Db CTCTAGTCAAAAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2316
QY ACAGATGATGATGCTGTGAAACAGTTTGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2405
Db ACAGATGATGATGCTGTGAG 2376
QY GGCAGGACATACCCAGGAG 2465
Db GGCCTGCAATGCGAG 2436
QY GTAGAAACAG 2525
Db GTAGAAACAG 2496
QY CTCTAGAAACCTCAAAACAG 2585
Db CTCTAGAAACCTCAATGAGTGAAG 2556
QY TATACATGAACTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2645
Db TATACATGAACTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2615
QY AGTTTCCACCAATTAATAATCCATGAGTAACTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2701
Db ATTTCCACCGGTTTAAATACATATATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2671

RESULT 13

MUSSEMAIII

LOCUS

DEFINITION

ACCESSION

VERSION

MUSSEMAIII 1998 bp mRNA linear ROD 03-AUG-1995
Mus musculus semaphorin III (Sema III) mRNA, partial cds.

L40484

L40484.1 GI:703189

Db 781 GCAACACATGCGAATTCGACAGCTATCAAGAATGATTTTGGTGGCCACAGAAGCTTA 840
Qy 1037 GTGATAAATGACAAACATTCCTCAAGCTCGTCTGATTTGCTCAGTGCAGGTCAAAAT 1096
Db 841 GTTAAACAAATGGACTACTTTCTTAAAGCTCGCTAAATTTGTTCACTGCTGCTGCTAAT 900
Qy 1097 GCAATTCACACTCATTTTCATGAACCTGAGGATGATTCCTCAATGAACCTTAAAGATCCT 1156
Db 901 GGAATTTGATACACATTTTATGATGAACCTACAGGATGTTGTTTATGATGAACCTCAAAAGATCCA 960
Qy 1157 AAAAAATCCAGTGTATATGAGAGTGTTCAGACTTCAGTAAACATTTTCAAGGGATCAGCC 1216
Db 961 AAAAAATCCAGTGTATATGAGAGTGTTCAGACTTCAGTAAACATTTTCAAGGGATCAGCA 1020
Qy 1217 GTGTGATGATAGATGATGATGAGAGGTTTCTTGGTCCATATGATGCCACAGG 1276
Db 1021 GTCTGTATGATGATGATGATGAGAGGTTTCTTGGTCCATATGATGCCACAGG 1080
Qy 1277 GATGACCCCAACTATCAATGGGTGCTTATCAAGGAAGTCCCTTATCCAGGCCAGGA 1336
Db 1081 GATGGTCCAAATTTATGATGAGGTTGCCATTCGAAGGAGGTTCCATACCTCGACTGGA 1140
Qy 1337 ACTGTGCCAGCAAAACATTTGGTGGTGTGATCTCTACAAAGGACCTTCTCGATGATGT 1396
Db 1141 ACTTGGCCCTAGGCAGACTTTTGGTGGTGTGATCTCTCAAAAGGATCTCCCTCGATGATGT 1200
Qy 1397 ATAACTTTGCAAGAGTCAATCCAGCCATGATCAATCCAGTGTTCCTATGAACATGCG 1456
Db 1201 ATTATGTTTGCAGAGTCAACCGGCCATGTCAATCTCTGTTTTCCTCAATCAATATCGT 1260
Qy 1457 CAAATAGTGCATCAAAACGAGTGAATTAATCAATTTACACAAATGTCGTAGACCGAGTG 1516
Db 1261 CCATCATTAATAAAGGAGTGGATTAATCAATTTTACAAATAGTGGTGGATCGAGTG 1320
Qy 1517 GATCGAAGAGTGGACAGTATGATGTATGTTTATCGGAACAGATGTTGGGACCGTCTT 1576
Db 1321 GAAGCTGAGGATGGTCAATACGATGTATGTTTATTTGGAACATGATAGGGGACCATTTTG 1380
Qy 1577 AAAGTAGTTCATTCCTTAAGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAAGAA 1636
Db 1381 AAAGTAGTCTCTGTTCCCAAGGAGACCTGGACTGATTTAGAGAGGTTCTTACTGGAAGAA 1440
Qy 1637 ATGACAGTGTTCGGAACCGACTGCTATTTTCAGCAATGGAGCTTCCACTAAGCAGCAA 1696
Db 1441 ATGACTGTATTTAGGAGGCAACTGCTATTTTCAAGCATGGAGATTTCAACAAAGCAACAA 1500
Qy 1697 CAACATATATTTGGTTCAACGGCTGGGGTGGCCAGCTCCCTTTTACACCGGTGTGATAT 1756
Db 1501 CAACATATATTTGGTCTGTCAGTGGGTGTTTCCCAACTCCCGCTTCCAGCGTGTGATGT 1560
Qy 1757 TAGGGAAAGGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTTGGGATGTT 1816
Db 1561 TATGAAAAGCTGTGCTGAGTGTTCCTTGGCCCTTGGCAGAGACCTTATTTGCTTGGGATGGA 1620
Qy 1817 TCTGATGTTCTGCTGCTATTTTCCACTGCAAGAGACGCAAGAGCACAAGATATAAGA 1876
Db 1621 TCTTATGCTACGCTACTTTTCTACTGCTTAAAGACGTACTAGGGCCCAAGATATCAGG 1680
Qy 1877 AATGGAGACCCACTGACTCTGCTTACACTTTTACACCAT- - -GATAATCACCATGGCCAC 1933
Db 1681 AAGGGGATCCCTTAACACATTTTTCAGACCTTACAGCATCAAGATGATCCACATAGACAA 1740
Qy 1934 AGCCCTGAAGAGAGATCATCTATGTTGATAGAGATAGTACACATTTTGAATGTCAGT 1993
Db 1741 AGCCTAGAGAAAGATATACTATGTTGTTGGAGAACAGTACAGTACATCTCTGGAATGCAGC 1800
Qy 1994 CGGAAGTCCAGAGAGCGGTGTTCTATTCGCAATTCAGAGCGCAAAATCAAGAGCGAAA 2053
Db 1801 CCTAATACAGCGTGCCTAGTTTCTGCGAGTTCCAGAGCAAAATGAAGAAAAG 1860
Qy 2054 GAAGAGATCAGAGTGGATCATATCATCAGGACAGATCAAGGCGCTTCTGCTACGTAGT 2113

Db 1861 GATGAGATAAAGTGTGATGAACGCAAAATAAAGACAGAACATGGCTTTCTTCTGCGCAGC 1920
Qy 2114 CTACACAGAGAGGATTCAGGCAATTTACCTCTGCCATGCGGTGGACATCGGTTTCATACAA 2173
Db 1921 CTAAAGAAAGAGAGATTCAGGCATATATTTATTTGCAATGCGATAGAGCATGGTTTTCATGAG 1980
Qy 2174 ACTCTTCTTAAGGTAACCTCTGGAAGTCAATGACACAGAGCATTTTGGAAAGAACTTCTTCAT 2233
Db 1981 ACCCTTCTCAAGTAACCTCTGGAGATATTTGATACAGAACATTTAGATGAGCTTCTTCAC 2040
Qy 2234 AAAGATGATGATGAGATGGCTCT- - -AAGACCAAGAAATGCTCAATAGCATGACACT 2290
Db 2041 AAAGAGATGAAGGAGGAGACAGTCACAAACACAAAGAACCTCAACAGATATGCTCCA 2100
Qy 2291 AGCCAGAGGTCCTGTCAGAGACTTTCATGAGCTCATCAACCCCACTCAACAGCTCAACAGC 2350
Db 2101 ACTCAAAAGATATGATAGAGACTTTTATGAGTTTAAACCCCAAAATCTTAAACACA 2160
Qy 2351 ATGATGATGTTCTGTGAACAAAGTTTGGAAAAGGACCGAAAACACACGTCGCGCAAGGCCA 2410
Db 2161 ATGATGATGTTTGTGAGCAGTGTGGAAGAGGACCGCAAGCAACGTCGCGCAAAATAAT 2220
Qy 2411 GGACATATCCCGAGGAAACAGTAAACAAATGGAAGACATTTACAGAAAATAAGAAAGGTAGA 2470
Db 2221 GGAATGTCCAGTTTAGCAACACAAATGGAAGCACCTTACAGGAAAACAAGAGGGGTAGA 2280
Qy 2471 AACAGGAGGCCACAGCAATTTGAGAGGGCACCCAGGAGTGTCTGA 2515
Db 2281 AATAGGAGGCCCATGAATTTGAGAGGGCACCAAGGAGTGTCTGA 2325

RESULT 15
LOCUS I11900 1481 bp DNA linear PAT 26-JUL-1995
DEFINITION Sequence 1 from Patent US 5416197.
ACCESSION I11900
VERSION I11900.1 GI:909343
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1481)
AUTHORS Raper,J.A. and Luo,Y.
TITLE Antibodies which bind human collapsin
JOURNAL Patent: US 5416197-A 1 16-MAY-1995;
FEATURES Location/Qualifiers
source
1..1481
/organism="unknown"
BASE COUNT 454 a 299 c 345 g 383 t
ORIGIN

Query Match 52.2%; Score 1415.2; DB 6; Length 1481;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1420; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

Qy 899 GAGAGTGACAACTCTGGAAGATGACAAAGTACTATTTTCTCCGTGAAAATGCAATAGAT 958
Db 53 GAACATGACAACTCTGGAAGATGACAAAGTACTATTTTCTCCGTGAAAATGCAATAGAT 112
Qy 959 GGAGAACACTCTGGAAGAGCTTACTCAGCTAGAAATAGGTTCAGATATGCAAGATGACTTT 1018
Db 113 GGAGAACACTCTGGAAGAGCTTACTCAGCTAGAAATAGGTTCAGATATGCAAGATGACTTT 172
Qy 1019 GGAGGGCACAGAGTCTGGTGAATAAATGGACAACTTCTCAAGCTCGTCTGATTTGC 1078
Db 173 GGAGGGCACAGAGTCTGGTGAATAAATGGACAACTTCTCAAGCTCGTCTGATTTGC 232
Qy 1079 TCAGTGCAGGTCCAAATGGCATTTGACACTCATTTTGGATGAATGACAGTGTATTCCTA 1138
Db 233 TCAGTGCAGGTCCAAATGGCATTTGACACTCATTTTGGATGAATGACAGTGTATTCCTA 292
Qy 1139 ATGAACCTTAAAGATCTTAAATAATCCAGTGTATATGAGTGTTTACAGCTTCAGTAAC 1198

Db 293 ATGAACCTTTAAAGATCCTAAAAATCCAGTTGTATATGGAGTGTATTACGACTTCCAGTAAC 352
QY 1199 ATTTTCAAGGATCAGCCGTGTGTATGTATAGCATGAGTGTGAGAAAGGTGTTCCTT 1258
Db 353 ATTTTCAAGGATCAGCCGTGTGTATGTATAGCATGAGTGTGAGAAAGGTGTTCCTT 412
QY 1259 GGTTCATATGCCACAGGATGAGCCCAACTATCAATGGTGCCTTATCAAGGAAGTTC 1318
Db 413 GGTTCATATGCCACAGGATGAGCCCAACTATCAATGGTGCCTTATCAAGGAAGTTC 472
QY 1319 CCTATCCACGCGCAGGACTTGTCCAGCAAAACATTTGGTGGTGTGACTCTCAAAAG 1378
Db 473 CCTATCCACGCGCAGGACTTGTCCAGCAAAACATTTGGTGGTGTGACTCTCAAAAG 532
QY 1379 GACCTTCCTGATGATGTTTATAACCTTTGCAAGAAGTCATCCAGCCATGTACAATCCAGTG 1438
Db 533 GACCTTCCTGATGATGTTTATAACCTTTGCAAGAAGTCATCCAGCCATGTACAATCCAGTG 592
QY 1439 TTTCTATGAACAATCGCCCAATAGTATCAAAACGGATGTAAATNTCAATTTACAA 1498
Db 593 TTTCTATGAACAATCGCCCAATAGTATCAAAACGGATGTAAATNTCAATTTACAA 652
QY 1499 ATTGTCGTAGACCGAGTGCAGAGATGACAGATGACAGATGATGTTTATCGGAACA 1558
Db 653 ATCGTCGTAGACCGAGTGCAGAGATGACAGATGATGTTTATCGGAACA 712
QY 1559 GATGTGGGACCGTTCTTAAAGTAGTTCATTTCTAAAGGACCTTGGTATGATTTAGAA 1618
Db 713 GATGTGGGACCGTTCTTAAAGTAGTTCATTTCTAAAGGACCTTGGTATGATTTAGAA 772
QY 1619 GAGTTCCTGCTGGAAGAAATGACAGTTCCTGGGAAACCGACTGCTATTTACGAATGGAG 1678
Db 773 GAGTTCCTGCTGGAAGAAATGACAGTTCCTGGGAAACCGACTGCTATTTACGAATGGAG 832
QY 1679 CTTTCCACTAAGCAGCAACAATATATGTTTCAACGGCTGGGGTTCGCCAGTCCCT 1738
Db 833 CTTTCCACTAAGCAGCAACAATATATGTTTCAACGGCTGGGGTTCGCCAGTCCCT 892
QY 1739 TTACACCGGTGTGATTTTACGGGAAAGCGTGTGCTGAGTGTGCTCGCCGAGACCCCT 1798
Db 893 TTACACCGGTGTGATTTTACGGGAAAGCGTGTGCTGAGTGTGCTCGCCGAGACCCCT 952
QY 1799 TACTGTCTTGGGATGGTTCGCAATGTTCTGCTATTTTCCACTGCAAAAGAGACGACA 1858
Db 953 TACTGTCTTGGGATGGTTCGCAATGTTCTGCTATTTTCCACTGCAAAAGAGACGACA 1012
QY 1859 AGACGACAGATATAAGAAATGGAGACCCACTGCTCACTGTTAGACTTACACCATGAT 1918
Db 1013 AGACGACAGATATAAGAAATGGAGACCCACTGCTCACTGTTAGACTTACACCATGAT 1072
QY 1919 AATCACCATGGCCACAGCCCTGAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTACACA 1978
Db 1073 AATCACCATGGCCACAGCCCTGAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTACACA 1132
QY 1979 TTTTGGAAATGAGTCCGAAGTCCGAGAGCGGTGCTATTTGGCAATTCAGAGCGGA 2038
Db 1133 TTTTGGAAATGAGTCCGAAGTCCGAGAGCGGTGCTATTTGGCAATTCAGAGCGGA 1192
QY 2039 AATGAAGAGCGAAAGAGAGATCAGAGTGGATGATCATATCATCAGCAGATCAAGGC 2098
Db 1193 AATGAAGAGCGAAAGAGAGATCAGAGTGGATGATCATATCATCAGCAGATCAAGGC 1252
QY 2099 CTTCTGCTACGTAGTCTACACAGAGAGATTCAGGCAATTTACCTCTGCGATCGCGTGGAA 2158
Db 1253 CTTCTGCTACGTAGTCTACACAGAGAGATTCAGGCAATTTACCTCTGCGATCGCGTGGAA 1312
QY 2159 CATGGGTTTACATAAACTCTTCTTAAGGTAAACCTCGAAGTCATTGACACAGACATTTG 2218
Db 1313 CATGGGTTTACATAAACTCTTCTTAAGGTAAACCTCGAAGTCATTGACACAGACATTTG 1372
QY 2219 GAAGAACTTCTTATGAAGATGATGAGATGGCTCTAAGACCAAAAGAAATGTCCAAT 2278
Db 1373 GAAGAACTTCTTATGAAGATGATGAGATGGCTCTAAGACCAAAAGAAATGTCCAAT 1432

QY 2279 AGCATGACACCTAGCCAGAGGTCCTGGTACAGAGACTTTCATGCAGCTC 2326
Db 1433 AGCATGACACCTAGCCAGAGGTCCTGGTACAGAGACTTTCATGCAGCTC 1480

Search completed: July 31, 2003, 15:53:21
Job time : 9588 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 12:47:35 ; Search time 5029 Seconds
(without alignments)
13092.230 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709

Sequence: 1 aatctttatttatcatgatg.....agggttttttctctaatacc 2709

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_estc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	723	26.7	723	12	BI870437 603395690
2	715.2	26.4	3317	11	AK052671 Mus muscu
3	715.2	26.4	3672	11	AK028900 Mus muscu
4	712	26.3	4248	11	AK031704 Mus muscu

5	674.8	24.9	679	12	BM711125
6	663.8	24.5	702	12	BM723807
7	620.2	22.9	2960	11	AK043634
8	620.2	22.9	3154	11	AK033597
9	620.2	22.9	3292	11	AK037034
10	617	22.8	3287	11	AK048364
11	594.6	21.9	729	14	CD349263
12	590.4	21.8	900	14	CB196368
13	588.6	21.7	889	10	BF667677
14	581	21.4	775	13	BU703431
15	567.4	20.9	699	14	CA327319
16	543.8	20.1	604	12	BI183160
17	537.8	19.9	773	13	BUS15916
18	524.2	19.4	3436	11	AK014333
19	522.4	19.3	3884	11	AK034239
20	510	18.8	570	10	BE033028
21	506.8	18.7	510	14	CD216784
22	504.2	18.6	3394	11	AK053632
23	501.8	18.5	698	13	BU461875
24	485.8	17.9	884	10	BF700780
25	475.4	17.5	559	14	CA872840
26	463.4	17.1	807	14	CB595903
27	458.6	16.9	2596	11	AK053115
28	433.2	16.7	660	10	BB612039
29	439.4	16.2	497	14	CA395261
30	428.2	15.8	731	13	BU708579
31	397.6	14.7	917	13	BQ884924
32	388.4	14.3	747	12	BQ018643
33	385.6	14.2	1809	11	AK051165
34	382	14.1	382	13	EX283116
35	378.4	14.0	493	14	CB725110
36	366.4	13.5	914	14	CA488964
37	364.6	13.5	787	14	CB518391
38	360.6	13.3	511	13	BU445100
39	353	13.0	441	14	CA871124
40	351.4	13.0	600	12	BI989239
41	344.4	12.7	721	2	BSM084323
42	341.4	12.6	692	13	BQ447821
43	336	12.4	839	10	BF667143
44	331.2	12.2	773	28	BH265538
45	325.2	12.0	394	14	CB776561

ALIGNMENTS

RESULT 1

BI870437

LOCUS

DEFINITION

603395690F1 NIH_MGC_90 Homo sapiens CDNA clone IMAGE:5405519 5', mRNA sequence.

ACCESSION

BI870437

VERSION

BI870437.1 GI:16044110

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

NIH-MGC http://mgc.nci.nih.gov/.

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE

Unpublished

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbe@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

CDNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Plate: L1AM12034 row: c column: 24

High quality sequence stop: 723.

BI870437 723 bp mRNA linear EST 11-OCT-2001
603395690F1 NIH_MGC_90 Homo sapiens CDNA clone IMAGE:5405519 5', mRNA sequence.

ACCESSION

BI870437

VERSION

BI870437.1 GI:16044110

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

NIH-MGC http://mgc.nci.nih.gov/.

AUTHORS

National Institutes of Health, Mammalian Gene Collection (MGC)

TITLE

Unpublished

JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbe@mail.nih.gov

Tissue Procurement: ATCC

CDNA Library Preparation: Life Technologies, Inc.

CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

CDNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Plate: L1AM12034 row: c column: 24

High quality sequence stop: 723.

FEATURES
source

Location/Qualifiers

1. .723
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5405519"
/issue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 90"
/note="Organ: liver; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC Library."
BASE COUNT 225 a 148 c 159 g 191 t
ORIGIN

Query Match 26.7%; Score 723; DB 12; Length 723;
Best Local Similarity 100.0%; Pred. No. 8.7e-173;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 454 CAAGATTTTCAAAAGATGTGCGCCAGTATCTTACACCGAAGAGATGATGCAAGTG 513
DB 1 CAAGATTTTCAAAAGATGTGCGCCAGTATCTTACACCGAAGAGATGATGCAAGTG 60
QY 514 GCGTGGAAAAGACATCCTGAAAAGATGTGCTAAATTTTCATCAAGGTACTTAAAGGCATATAA 573
DB 61 GCGTGGAAAAGACATCCTGAAAAGATGTGCTAAATTTTCATCAAGGTACTTAAAGGCATATAA 120
QY 574 TCAGACTCACTGTGACGCTGTGGAACCGGGGCTTTTCATCAATTTTCACCTACATTTGA 633
DB 121 TCAGACTCACTGTGACGCTGTGGAACCGGGGCTTTTCATCAATTTTCACCTACATTTGA 180
QY 634 AATTCGACATCATCTGAGGCAATATTTTAAAGTGTGAGAACTCACATTTTGAACCGG 693
DB 181 AATTCGACATCATCTGAGGCAATATTTTAAAGTGTGAGAACTCACATTTTGAACCGG 240
QY 694 CGGTGGGAAGAGTCCATATGACCTTAAGCTGTGACAGCATCCCTTTTAAATAGATGGAGA 753
DB 241 CGGTGGGAAGAGTCCATATGACCTTAAGCTGTGACAGCATCCCTTTTAAATAGATGGAGA 300
QY 754 ATTATCTCTGAACTGACGCTGATTTTATGGGGGAGACTTTGCTATCTTCCGAACCTCT 813
DB 301 ATTATCTCTGAACTGACGCTGATTTTATGGGGGAGACTTTGCTATCTTCCGAACCTCT 360
QY 814 TGGGCAACCAACCAATCAGACAGCAGCATGATTCAGGTGGCTCAATGATCCAA 873
DB 361 TGGGCAACCAACCAATCAGACAGCAGCATGATTCAGGTGGCTCAATGATCCAA 420
QY 874 GTTCATTAGTGGCCACCTCATCTCAGAGAGTGACAAATCCTGAAGATGACAAAGTATACTT 933
DB 421 GTTCATTAGTGGCCACCTCATCTCAGAGAGTGACAAATCCTGAAGATGACAAAGTATACTT 480
QY 934 TTTCTTCGTTGAAATGCAATAGATGAGAGAACACTCTGGAAGACTACTACGCTAGAT 993
DB 481 TTTCTTCGTTGAAATGCAATAGATGAGAGAACACTCTGGAAGACTACTACGCTAGAT 540
QY 994 AGGTGAGATTCAGAGATGACTTTGGAGGGGACAGAGATCTGTGATGATGATGACAAAC 1053
DB 541 AGGTGAGATTCAGAGATGACTTTGGAGGGGACAGAGATCTGTGATGATGATGACAAAC 600
QY 1054 ATTCTCTCAAGAGTCTGCTGATTTGCTCAGTGCAGGTCCAAATGCGATTCGACCTCATTT 1113
DB 601 ATTCTCTCAAGAGTCTGCTGATTTGCTCAGTGCAGGTCCAAATGCGATTCGACCTCATTT 660
QY 1114 TGATGAATCGAGGATGATTCCTAATGAACTTTAAAGATCCTAAAAATCCAGTTGTATA 1173
DB 661 TGATGAATCGAGGATGATTCCTAATGAACTTTAAAGATCCTAAAAATCCAGTTGTATA 720
QY 1174 TGG 1176
DB 721 TGG 723

RESULT 2
AK052671
LOCUS
DEFINITION

AK052671 3317 bp mRNA linear HTC 05-DEC-2002
Mus musculus 0 day neonate kidney cDNA, RIKEN full-length enriched
library, clone:D630018P03 product:SMAPHORIN 3D PRECURSOR homolog
[Homo sapiens], full insert sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AK052671.1 GI:26342870
HTC; CAP trapper.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

REFERENCE
AUTHORS

1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636

REFERENCE
AUTHORS

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

REFERENCE
AUTHORS

3 Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Ozawa, K., Tanaka, T., Ohara, E., Watanabe, J.,
Yoneda, Y., Ishikawa, T., Tanaka, K., Tanaka, T., Matsuo, S., Kawai, M.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

TITLE

JOURNAL
MEDLINE
PUBMEDREFERENCE
AUTHORS

4 Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadoya, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,
Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G.,
Quackenbush, J., Schriml, L.M., Staudli, F., Suzuki, R., Tomita, M.,
Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H.,
Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N.,
Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C.,
Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D.,
Hofmann, M., Hume, D.A., Kamiya, M., Lee, N.H., Lyons, P.,
Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P.,
Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H.,
Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H.,
Toyo-oka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L.,
Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S.
and Hayashizaki, Y.

Functional annotation of a full-length mouse cDNA collection
Nature 409 (6821), 685-690 (2001)
21085660
11217851

REFERENCE
AUTHORS

5 The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

REFERENCE
AUTHORS

6 (bases 1 to 3317)
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,

Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kaga, I., Kaskawa, T., Kato, H., Kawaj, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Satoh, H., Shibata, K., Shingawa, A., Shiraki, T., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.go.jp, FAX: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: <http://genome.gsc.riken.go.jp/>

URL: <http://fantom.gsc.riken.go.jp/>

source

source

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:D630018P03"
/db_xref="taxon:10090"
/clone="D630018P03"
/tissue_type="kidney"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
548..2881
/note="unnamed protein product; SEMAPHORIN 3D PRECURSOR
homolog [Homo sapiens] (SWISSPROT|O95025, evidence: FASTA,
92.8%ID, 100%length, match=2331)
putative"

```

```

/codon_start=1
/protein_id="BA35092.1"
/db_xref="GI:26342871"
/translation="MNYVDENPFRSSQDLHLFHAMWIMTWLFLVPVTTSTKONIPR
LKITYDILLNSNTICPLGSGEGLDFTLLIDBERGILLKAKDHFVLLSLVDLNNKFN
KTYLPAAKERVELCKLAGKDANAECANFIRLVQPNKTHVYVCGTCAFAHPLCGVIDL
GANKELIFKLDTHNLESGRLKCPFPQOPQFPASVWTDHLYLSCPTASDFLGKDTAFTSR
LGLMQDHSIRTDISEHWWLKGPGITGTFIDTNPDDDKIYFFPFRSSQEGSTDR
SILSRVGVVGNDDGGORSLLKNWTKTLKARLCSIPGSCADTHDELQIYLLPFTY
DERNPVGVYFTTTSIFGSAVCVSMADIRAVNFGPYAHKSDASHRWYQDQIRPY
PREGTCPTSTYDLIKSTDPDDVIFSIIRHPVYKSVYVAGAPTFKRIINVYRLT
QIVVDHVAEDGQYDWMFLGTDLGTVLKVVISKEKNWEEVVLLELQVFKHPHTAIL
MELSLKQQLIYVSGWGLVQILSHRCDTYGKACADCCCLARDPYCWDGNACSVAPTS
KRARRQDKIGDPIQCDWEDVISHETADEKVI FGIEPNSTFLECIPKSKQAQASVW
YIORSGDHEHEELKPERIITKDYLGLLSRLKQDGMWYKQAQEHFTHTIKTLIN
VINEQMENTQRAEYQEGQVKOLLARSRLRYKQYIQLLSPNFSLDQYCEQWMYKEKR
RQRNKSPKWKHMQEKNKKRNRHRLDELQRSVAT"
939 a 756 c 807 g 815 t

```

722	ATCCCCTTTTTGGGTTCATCAGAAGGATTTGGATTTCCAGACTCTTCTTTTGGATGAGGAG	781
395	CGGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATATATTTTCATTGCACTGGTTAA---T	451
782	AGGGCATACTGCTCTTAGAGGCCAAGACCATGTCTTCTGCTCAGTCTGGTTGACTTG	841
452	ATCAAGGATTTTCAAAGATTTGTGTGGCCAGTATCTTTACACCAAGAGAGATGAATGCAAG	511
842	AACAAGAAATTTTAAGAAGATTTATGCGCTGCTGCAAAAGAACGAGTGGAGCTATGTAAA	901
512	TGGCTGGAAAAGACATCTCTGAAGAATGTGCTAAATTTTCATCAAGGTACTTTAAGGCATAT	571
902	TTAGCTGGGAAAGATGCGCAATGCAGAAATGTGCAAAATTTTCATCGTGTGCTTCAACCCAT	961
572	AATCAGACTCACATTGTACGCCTCTGGAAACGGGGGCTTTTCATCCAAATTTTGCACCTACATT	631
962	AATAAGACTCACGTTTACGTGTGTGGNACTGGAGCGTTTTCATCCGCTGTGTGGGTACATT	1021
632	GAATTTGGACATCATCTGTAGGACAATATTTTAAAGCTGGAGAACTCACATTTTGTGAAAAC	691
1022	GATCTCGGCCCAACAAGGAGGAACATATATTTTAAACTAGACACGCAACCTGGAGTCT	1081
692	GGCGTGGGAAGAGTCCATATGACCCCTAAGCTCTGCACAGCATCCCTTTTAATAGATGGA	751
1082	GGCAGACTGAATGTCTCCCTTGTATCTCTCAGCAGCCCTTTTGCCTTTCAGTAATGACAGATGAG	1141
752	GAATTTACTCTCGAACTGCAGCTGTATTTTATGGGCGAGACTTGTCTATCTTCCGAACCT	811
1142	CACCTCTACTCTGGACAGCTTCTGATTTTCCCTTGGCAAGACACTGCAATTCACAAAGTCT	1201
812	CTTGGGC-----ACCACCAACCAATCAGACAGAGCAGCATGATTTCCAGGTGGCTC	862
1202	CTGGGGCTAATCAGGACCACCACTTCATCAGAACTGCATATTCAGAGCACCCTCGCTC	1261
863	AATGATCCAAAGTTCATTAGTGGCCACCTCATCTCAGAGAGTGACAACTCCTGAAGATGAC	922
1262	AATGGAGCAAAATTTATCGGAACTTTCCCAATTTCCAGACACCTATTAATCCAGATGATGAT	1321
923	AAAGTATACTTTTCTTCGTGAAATGCAATAGATGGAGAACACTCTGGAAGACTACT	982
1322	AAATATATTTCTTCTTTTGAGAAATCATCCAGGAAGGCAGTACTTCTGACAGAGCAATT	1381
983	CAGCTAGAATAGGTCCAGATATGCAAGAATGACTTTGGAGGGCAGAGAAGTCTGTGTAAT	1042
1382	CTTTCAAGAGTTGGAAGAGTTGTTAAGATGATGATAGTGGGCAACGAAGTCTGATAAAC	1441
1043	AAATGGACAACTTCTCAAAGTCTGCTGATTTGCTCAGTGGCCAGGTCCAAATGCAATT	1102
1442	AAATGGACAACTTTTCTAAAGGCAAGACTGATTTGCTCGATTTCTCTGGAAGCGATGGGCA	1501
1103	GACACTCATTTTGTATCAACTGCAAGATGATTTCTTAATGAACTTTTAAAGATCCTAAAAAT	1162
1502	GATACCCATTTTATGAACTCCAAAGATATTTACTTACTCCCTACGAGAGATGAAGAAAT	1561
1163	CCAGTTGTATGAGAGTGTTTACGACTTCAGGTAAATTTTCAAAGGGAATCAGCCGTGTGT	1222
1562	CCGTGTAGTATATGGAGTCTTTACCAACAACGAGTCCATCTTCAAAGGCTCTGCTGTCTGT	1621
1223	ATGTTATAGATGATGATGTGAGAGGGTGTTCCTTTGGTCCATATGCCCCACAGGGATGGA	1282
1622	GTGTACAGCATGCTCATATCCGAGCAGTCTTTTAAATGGTCCCTATGCTCATGAAGAAAGT	1681
1283	CCCAACTATCAATGGGTGCTTTATCAAGGAAGAGTCCCCATATCCAGGGCCAGGAATTCTGT	1342
1682	GCTGACCATCGCTGGGTGCAATATGATGAAGAGATACCTTTACCCCGCACTCGAAGCGTGT	1741
1343	CCGAGCAAAACATTTG---GTGCTTTTGTACTCTACAAAGGAGCTTCTCTGATGATTTATA	1399
1742	CCAAGCAAAACCTATGACCCCATGATTAAGTCCACCCGAGACTTTTCCAGACGATGTTATC	1801
1400	ACCTTTGCAAGAGTTCATCCAGCCATGTACAATCCAGTGTTTTCCCTATGAACAATCGCCA	1459
1802	AGTTTTCATTAAGCGGGCACCCTGTGATGATTAAGTCCGTGTATCCCACTGGCAGGACCG	1861

Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN,
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>

FEATURES

source

location/Qualifiers
1..3672
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:4732468K20"
/db_xref="taxon:10090"
/clone="4732468K20"
/tissue_type="skin"
/dev_stage="10 days neonate"
/note="unnamed protein product; SEMAPHORIN 3D PRECURSOR
homolog [Homo sapiens] (SWISSPROT|095025, evidence: FASTA,
92.8%ID, 100%length, match=2331)
putative"
/codon_start=1
/protein_id="BAC26185.1"
/db_xref="GI:26324862"
/translation="MNVTKDENPRSRQDLHFLHAWMLIMTVLFLPVTETSKQNI
KLIYPAKSERVELCKLAGDANAEANFIRVLOPKNTKVVYVCGTGAFLPLCGYIDL
GANKBELIFKLDTHNLSGRLLKCPFPQPPASVMTDEHLVSGTASPLGKDAFTSR
LGLMDHHSIKTDISEHNLNGAKFIGTPIIDTINPDNDKDIYFFRESQSGSTSR
SILSRGVKNDVQGRSLINKWTFKRLKLSIPSGDGAHDFDELQDIYLLPTR
DERNPVYGVFTTSSIFKSAVCVYSNADIRAVFNGPYAHKESADHRWQYDGRIPY
PRPGTSPSKTDLPIKSTPPDPDVISFIRRHVPMVKSVPVAGAPFKRINVDYRLT
QIVDVHVARDQYDVMFLGDTIGTVLKVSIKSKNMWEVLEELQVFKHPITALIN
MELSILKQQLVGVSDGLVQLSLRCDTYGKACDCLARDPYCAMDGNACSRVAPIS
XRRARQDKVTGDPITQCDWIEDISHEADEKVFGLFEFTSTFLECIPKSGQASVW
VIENQEMHREBELPDERIKIDYGLLRLSLQKDSGMYCKAQEHTFIHTIVKLTLN
VIENQEMHREBELPDERIKIDYGLLRLSLQKDSGMYCKAQEHTFIHTIVKLTLN
RQRNKGSPKWKHMOMKRRNRHRLDELQRSVAT"

CDS

329..3662
/note="unnamed protein product; SEMAPHORIN 3D PRECURSOR
homolog [Homo sapiens] (SWISSPROT|095025, evidence: FASTA,
92.8%ID, 100%length, match=2331)
putative"
/codon_start=1
/protein_id="BAC26185.1"
/db_xref="GI:26324862"
/translation="MNVTKDENPRSRQDLHFLHAWMLIMTVLFLPVTETSKQNI
KLIYPAKSERVELCKLAGDANAEANFIRVLOPKNTKVVYVCGTGAFLPLCGYIDL
GANKBELIFKLDTHNLSGRLLKCPFPQPPASVMTDEHLVSGTASPLGKDAFTSR
LGLMDHHSIKTDISEHNLNGAKFIGTPIIDTINPDNDKDIYFFRESQSGSTSR
SILSRGVKNDVQGRSLINKWTFKRLKLSIPSGDGAHDFDELQDIYLLPTR
DERNPVYGVFTTSSIFKSAVCVYSNADIRAVFNGPYAHKESADHRWQYDGRIPY
PRPGTSPSKTDLPIKSTPPDPDVISFIRRHVPMVKSVPVAGAPFKRINVDYRLT
QIVDVHVARDQYDVMFLGDTIGTVLKVSIKSKNMWEVLEELQVFKHPITALIN
MELSILKQQLVGVSDGLVQLSLRCDTYGKACDCLARDPYCAMDGNACSRVAPIS
XRRARQDKVTGDPITQCDWIEDISHEADEKVFGLFEFTSTFLECIPKSGQASVW
VIENQEMHREBELPDERIKIDYGLLRLSLQKDSGMYCKAQEHTFIHTIVKLTLN
VIENQEMHREBELPDERIKIDYGLLRLSLQKDSGMYCKAQEHTFIHTIVKLTLN
RQRNKGSPKWKHMOMKRRNRHRLDELQRSVAT"

BASE COUNT

1091 a 787 c 855 g 939 t

ORIGIN

Query Match 26.4%; Score 715.2; DB 11; Length 3672;
Best Local Similarity 62.1%; Pred. No. 1.7e-170;
Matches 1220; Conservative 0; Mismatches 723; Indels 21; Gaps 5;
275 AAGACATGTGCCAAGCTGAATATTCCTACAAAGAAATGTTGGAATCCAAACATGTG 334
443 AAACAAATATTCAAAGACTCAAGCTACCTACAAAGACTGCTGCTTTCAACACCTGT 502
335 ATCACTTTCAATGGCTTGGCCAAAGCTCCAGTTATCATCTTCTTTCGATGAGGAA 394
503 ATCCCTTTTGGGTTTATCAGAGGATGAGTTCCAGACTCTTCTTTTGGATGAGGAG 562
395 CGGAGTAGGCTGTATGTTGGAGCAAGGATCACATATTTTCATTCGACCTGGTTAA--T 451
563 AGGGGCACTGCTCTAGGAGCCAAAGACCACTGTCTTCTGCTCAGTCTGTTGACTTG 622
452 ATCAAGGATTTTCAAGATTTGTGTGGCAGTATCTTACACAGAGAGATGAATGCAAG 511
623 AACAGGATTTTAAAGAGATTTATTTGGCTCTCTGCAAAAGAACGAGTGGAGCTATGAAA 682
512 TGGGCTGGAAGACATCTTCAAGAGATGCTATTTTCATCAAGGATCTTAAGGCATAT 571
683 TTACTGGGAAGATGCCAATGCAAGATGTGCAATTTTCATCCGTGTGCTTCAACCCCTAT 742
572 AATCAGACTCACTGTGTAGCCCTGTGGACCGGGGCTTTTCATCCAAATTTGCACCTACAT 631
743 AATRAAGACTCAGTTTACGTGTGTGGAAGTGGAGCGTTTCATCCGCTGTGTGGGTACAT 802
632 GAAATTGACATCATCTCTGAGGACAATATTTTAAAGCTGGAGAACTCACATTTTGAAGAC 691

803 GATCTCGGCGCCAAACAGGAGAACTCATATTTTAACTAGACACGCAACCTGGAGTCT 862
692 GGCCGTGGGAGAGTCCATATAGCCTTAAGCTGTGACAGCATCCCTTTTAAATAGATGA 751
863 GGCAGACTGAAATGCTCCCTTTGATCCTCAGCAGCCTTTTGTCTCAGTAATGACAGATG 922
752 GAATATATCTGGAAGTGCAGCTGATTTTATGGGGGAGAGACTTTTGTCTATCTTCGAA 811
923 CACCTCTACTCTGGAACAGCTTCTGATTTCTTGGCAAGAGACACTGCTATTCACAAG 982
812 CTTGGGC-----ACCACCCCAATCAGGACAGCAGCAGCATGATTCACAGGTGGCTC 862
983 CTGGGGCTAATGACAGGACCACTTCCATCAGAACTGACATTTAGAGCACCACCTGGCTC 1042
863 AATGATCCAAAGTTTCATTAGTGGCCACCTCATCTCAGAGAGTGACAACTCTCGAAGAT 922
1043 AATGGAGCAAAATTTATCGGAACATTTCCCATTCAGACACCTATATATCCAGATGAT 1102
923 AAAGTATATCTTTCTTCCGTGAAATGCAATAGATGAGAGAACACTCTGGAAGAGCTACT 982
1103 AAAATATATTTCTTCTTTCGAGAAATCATCCAGGAAGGACGACTACTTCTGACAGAAG 1162
983 CACGCTAGATAGTGCAGATATGCAAGAACTGCTTTGGAGGGGACAGAGTCTGCTGAAT 1042
1163 CTTTCAAGAGTTGGAAGAGTTTGTAAAGATGATGATGTTGGGGCAACGAAAGTCTGATA 1222
1043 AAATGGACAACATTTCTCAAGAGTCTGCTGATTTTGTCTCAGTCCAGGTCCTCAATGG 1102
1223 AAATGGACAACATTTCTTAAGGCAAGACTGATTTTGTCTGATTTCTTGGAGCGATGG 1282
1103 GACACTCATTTTGTAGTAATGTCAGGATGATTTCTTAATGAATTTTAAAGATCTTAA 1162
1283 GATACCAATTTTGTAGTAATGTCAGGATGATTTTCTTACTTCTTCCCTAGCAGAGAT 1342
1163 CCAGTTCTATAGTGGTGTTCAGACTCCAGTAACTTTTCAAGGATCAGCGCTGTGT 1222
1343 CCTGTAGTATAGGAGTCTTTTACCAACACGACTCCATCTTCAAGGCTCTGCTGTCTGT 1402
1223 ATGTATAGCATGAGTGTGATGAGAGAGGTGTTCTTGTCTCATATGCCCACAGAGATGA 1282
1403 GTGTACAGATGCTGATATCCGAGCAGTCTTTAATGTTCTTATGCTCATAGGAAGT 1462
1283 CCCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTTATCCAGGCCAGGAATTTGT 1342
1463 GCTGACCATCGCTGGGTGCAATATGATGAAGAGGATACCTTACCCCGACCTGGAACG 1522
1343 CCAGCAGCAAACTTTG---GTGGTTTTGACTCTTCAAGAGGACCTTCTCTGATGATTTATA 1399
1523 CCAGCAGCAAACTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1582
1400 ACCTTTTCAAGAGAGTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1459
1583 AGTTTCAATAGGCGGACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1642
1460 ATAGTGTCAAAACGAGTGAATTTTCAATTTTCAAAATTTTCAAAATTTTCAAAATTTT 1519
1643 ACCTTCAAGAGATCAACGTTGATTTACAGCTGACGAGATAGTGGTGGATCACTGCTG 1702
1520 GCAGAGATGACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1579
1703 GCTGAAGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1762
1580 GTAGTTTCAATTTCTCAAGGAGACTTGTGATGATGATGATGATGATGATGATGATGATGAT 1639
1763 GTTGTGAGCATCTCAAGGAGAGAGTGG---AATATGAAGAGGTCGTAAGGAGAGCTT 1819
1640 ACAGTTTTTTCGGGAACCGACTGCTATTTTCAGCAATGAGGCTTTTCTCACTAAGCAG 1699
1820 CAGGATTTCAAGCACCACCAACAGCTATCTTGAACATGAGTGTGCTGTAAGAGCAGCA 1879
1700 CTATATATTCGTTCAACGGCTGGGTGGCCAGCTCCCTTTTACACCGGTGTGATTTTAC 1759
1880 TTGTACGTTGTTCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1939

Db 2007 TCCCAACAGCCTCTGTGGTGGTACATCCAGCGGTGAGGATGAGCATCGAGGAG 2066
 QY 2060 ATCAGATGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGTGTCTACAA 2119
 Db 2067 TTGAACCTGATGAAGGATCATCAAACTGACTATGGCTACTGATTCGAAGTCTGCAG 2126
 QY 2120 CAGAAGGATTCAGGCAATPACCTCTGCCATGCGGTGGAACTATGGGTTCATACAAACTCTT 2179
 Db 2127 AAGAAGGATTCGGGATGTATTACTGCAAGGACACAGGAGCACATTTTCATCCACACCATA 2186
 QY 2180 CTTAAGGTAACTGGAAGTCTATTGACACAGAGCAATTTGGAAGA 2223
 Db 2187 GTGAAGCTGACTTTGAATGTCTATTGAGATGAACAGATGGAAGA 2230

RESULT 5

BM711125 679 bp mRNA linear EST 28-FEB-2002
 LOCUS UI-E-DX1-aggv-1-12-0-UI.r1 UI-E-DX1 Homo sapiens cDNA clone
 DEFINITION UI-E-DX1-aggv-1-12-0-UI 5', mRNA sequence.

ACCESSION BM711125

VERSION BM711125.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 679)
 Ronaldo,M.F., Lennon,G. and Soares,M.B.

AUTHORS

Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL

MEDLINE

PUBMED

COMMENT

Genome Res. 6 (9), 791-806 (1996)

97044477

889548

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 Reverse.

Location/Qualifiers

1..679

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-E-DX1-aggv-1-12-0-UI"

/tissue_type="fetal eyes"

/dev_stage="fetal"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone_lib="UI-E-DX1"

/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

UI-E-DX1 is a normalized cDNA library containing the

following tissue(s): fetal eyes. The library was

constructed according to Ronaldo, Lennon and Soares,

Genome Research, 6:791-806, 1996. First strand cDNA

synthesis was primed with an oligo-dr primer containing a

Not I site. Double stranded cDNA was ligated to an EcoR I

adaptor, digested with Not I, and cloned directionally

into pT73-Pac vector. The oligonucleotide used to prime

the synthesis of first-strand cDNA contains a library tag

sequence that is located between the Not I site and the

(dT)18 tail. The sequence tag for this library is

AGATCAAGA. This library was created for the program, Gene

Discovery in the Visual System, supported by National Eye

BASE COUNT 242 a 133 c 161 g 142 t 1 others
 ORIGIN Institute (NEI)."

Query Match 24.9%; Score 674.8; DB 12; Length 679;
 Best Local Similarity 99.6%; Pred. No. 1.6e-160;
 Matches 676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1964 GAGATAGTAGCACATATTTTGGATGAGTCCGAGTCCGAGAGCGCTGCTCTATTGG 2023

Db 1 GAGATAGTAGCACATATTTTGGATGAGTCCGAGTCCGAGAGCGCTGCTCTATTGG 60

QY 2024 CAATTTCCAGAGCGCAATTAAGAGCGAAAGAGAGATCAGAGTGGATGATCATATCATC 2083

Db 61 CAATTTCCAGAGCGCAATTAAGAGCGAAAGAGAGATCAGAGTGGATGATCATATCATC 120

QY 2084 AGGACAGATCAAGCCCTTCTGCTAGTGTCTTCAACAGAGAGATTCAGGCAATTACCTC 2143

Db 121 AGGACAGATCAAGCCCTTCTGCTAGTGTCTTCAACAGAGAGATTCAGGCAATTACCTC 180

QY 2144 TGCCATGCGGTGGAACTGGTTTCATACAACTCTTCTTAAGGTAAACCTCGAAAGTCAAT 2203

Db 181 TGCCATGCGGTGGAACTGGTTTCATACAACTCTTCTTAAGGTAAACCTCGAAAGTCAAT 240

QY 2204 GACACAGAGCATTTGGAGAACTTCTTTCATAAAGATGATGAGGATGGCTCTAAGACC 2263

Db 241 GACACAGAGCATTTGGAGAACTTCTTTCATAAAGATGATGAGGATGGCTCTAAGACC 300

QY 2264 AAAGAAATGTCGAATAGCATGACACCTAGCCAGAAAGTCTGGTACAGAGATTCATGCAG 2323

Db 301 AAAGAAATGTCGAATAGCATGACACCTAGCCAGAAAGTCTGGTACAGAGATTCATGCAG 360

QY 2324 CTGATCAACCAACCCCAATCTCAACAGCATGATGAGTCTGTGTAACAAAGTTTGAAAGG 2383

Db 361 CTGATCAACCAACCCCAATCTCAACAGCATGATGAGTCTGTGTAACAAAGTTTGAAAGG 420

QY 2384 GACCGAAAAACAACGTCGGCAAGGCCAGGACATACCCAGGAGAACAGTAAACAATGGAAG 2443

Db 421 GACCGAAAAACAACGTCGGCAAGGCCAGGACATACCCAGGAGAACAGTAAACAATGGAAG 480

QY 2444 CACTTACAGAAATAAGAAAGGTAGAAACAGGAGGACCCACGAAATTTGAGAGGCGACCC 2503

Db 481 CACTTACAGAAATAAGAAAGGTAGAAACAGGAGGACCCACGAAATTTGAGAGGCGACCC 540

QY 2504 AGGAGTCTCTGAGCTGATTAACCTCTAGAAACCTCAACCAAGTAGAAACTTGCCTAGACA 2563

Db 541 AGGAGTCTCTGAGCTGATTAACCTCTAGAAACCTCAACCAAGTAGAAACTTGCCTAGACA 600

QY 2564 ATAACTCGAAAAACAATGCAATATATACATGAACCTTTTTCATGGCATTATGCGATGTTT 2623

Db 601 ATAACTCGAAAAACAATGCAATATATACATGAACCTTTTTCATGGCATTATGCGATGTTT 660

QY 2624 ACAATGTTGGAAATTCAG 2642

Db 661 ACAATGTTGGAAATTCAG 679

RESULT 6

BM723807

LOCUS 702 bp mRNA linear EST 01-MAR-2002

DEFINITION UI-E-E01-aiX-o-04-0-UI.r1 UI-E-E01 Homo sapiens cDNA clone

UI-E-E01-aiX-o-04-0-UI 5', mRNA sequence.

ACCESSION BM723807

VERSION BM723807.1

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 702)

Ronaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

889548

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 Reverse.

Location/Qualifiers

1..679

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="UI-E-E01-aiX-o-04-0-UI"

/tissue_type="fetal eyes"

/dev_stage="fetal"

/lab_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone_lib="UI-E-E01"

/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

UI-E-E01 is a normalized cDNA library containing the

following tissue(s): fetal eyes. The library was

constructed according to Ronaldo, Lennon and Soares,

Genome Research, 6:791-806, 1996. First strand cDNA

synthesis was primed with an oligo-dr primer containing a

Not I site. Double stranded cDNA was ligated to an EcoR I

adaptor, digested with Not I, and cloned directionally

into pT73-Pac vector. The oligonucleotide used to prime

the synthesis of first-strand cDNA contains a library tag

sequence that is located between the Not I site and the

(dT)18 tail. The sequence tag for this library is

AGATCAAGA. This library was created for the program, Gene

Discovery in the Visual System, supported by National Eye

JOURNAL
MEDLINE
PUBMED
COMMENT

Genome Res. 6 (9), 791-806 (1996)
9704477
889548
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Reverse.

FEATURES

Location/Qualifiers
1..702
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-E01-aix-o-04-0-UI"
/tissue_type="fetal eye"
/dev_stages="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-E01"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-E01 is a normalized cDNA library containing the
following tissue(s): fetal eye. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
CGCTATACC. This library was created for the program, Gene
Discovery in the Visual System, supported by National Eye
Institute (NEI)."

BASE COUNT 248 a 145 c 164 g 145 t

Query Match 24.5%; Score 663.8; DB 12; Length 702;
Best Local Similarity 99.3%; Pred. No. 1e-157;
Matches 698; Conservative 0; Mismatches 2; Indels 3; Gaps 3;

Qy 1920 ATCCATGGCCACAGCCCTGAAGAGAGATCATCTATGTTGTAGAGAAATAGTACACAT 1979
Db 1 ATCCATGGCCACAGCCCTGAAGAGAGATCATCTATGTTGTAGAGAAATAGTACACAT 59

Qy 1980 TTTTGGATGCGTCCGAAGTCCGAGAGAGCGTGGTCTATTGGCAATTCAGAGGCGAA 2039
Db 60 TTTTGGATGCGTCCGAAGTCCGAGAGAGCGTGGTCTATTGGCAATTCAGAGGCGAA 119

Qy 2040 ATGAAGCGCAAGAGAGATCAGTGGTATCATATCATCAGGACATCAAGGCC 2099
Db 120 ATGAAGCGCAAGAGAGATCAGTGGTATCATATCATCAGGACATCAAGGCC 179

Qy 2100 TTCTGCTACGTAGTCTACAACAGAGGATTTCAGGCAATTTACTCTGCGCATCGGTGGAA 2159
Db 180 TTCTGCTACGTAGTCTACAACAGAGGATTTCAGGCAATTTACTCTGCGCATCGGTGGAA 239

Qy 2160 ATGGGTTCAACAACTCTTTAAGGTAACCCCTGGAAGTCATTGACACAGAGATTGG 2219
Db 240 ATGGGTTCAACAACTCTTTAAGGTAACCCCTGGAAGTCATTGACACAGAGATTGG 299

Qy 2220 AAGAACTCTTCAATAAGATCATGATGAGATGCTCTTAAGACCAAGAAATGTCATA 2279
Db 300 AAGAACTCTTCAATAAGATCATGATGAGATGCTCTTAAGACCAAGAAATGTCATA 359

Qy 2280 GCATGACACCTAGCCAGAGGTTCTGGTACAGAGACTTTCATGACGCTCATCAACACCCCA 2339
Db 360 GCATGACACCTAGCCAGAGGTTCTGGTACAGAGACTTTCATGACGCTCATCAACACCCCA 419

Qy 2340 ATCTCAACACGATGATGATGTTCTGTGAACAGTTTGGAAAAGGACCGAAAACAGCTC 2399
Db 420 ATCTCAACACGATGATGATGTTCTGTGAACAGTTTGGAAAAGGACCGAAAACAGCTC 479

Qy 2400 GGCAAGCGCAGGACATACCCAGGGAACAGTAACTAAGTGAAGCCTTCAAGAAAATA 2459
Db 480 GGCAAGCGCAGGACATACCCAGGGAACAGTAACTAAGTGAAGCCTTCAAGAAAATA 539

Qy 2460 AGAAAGGTAGAAAACAGGAGGCCACCGAA-TTTGAGAGGGCCACCGAGGTGTCTGAGCT 2518
Db 540 AGAAAGGTAGAAAACAGGAGGCCACCGAA-TTTGAGAGGGCCACCGAGGTGTCTGAGCT 599

Qy 2519 GCATTACTCTAGAAAACCTCAACAGTAGAAAACCTTGCCTAGACAAATTAACGAAAACA 2578
Db 600 GCATTACTCTAGAAAACCTCAACAGTAGAAAACCTTGCCTAGACAAATTAACGAAAACA 659

Qy 2579 AATCAATATACATGAAC-TTTTTTCATGGCATTATGTGGATG 2620
Db 660 AATCAATATACATGAAC-TTTTTTCATGGCATTATGTGGATG 702

RESULT 7
AK043634
LOCUS
DEFINITION
Mus musculus 10 days neonate cortex cDNA, RIKEN full-length
enriched library, clone:AB30014F15 product:SEMAPHORIN 3B PRECURSOR
(SEMAPHORIN A) (SEMA A), full insert sequence.
AK043634
VERSION AK043634.1 GI:26089859
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus
ORGANISM Mus musculus

REFERENCE
1
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636

REFERENCE
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159

REFERENCE
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, K.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

REFERENCE
4
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y.,
Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S.,
Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I.,
Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R.,
Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T.,
Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H.,

RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama Kanagawa 230-0045, Japan (E-mail:genome-res@gs.riken.go.jp, URL:http://genome.gsc.riken.go.jp/, Tel.:81-45-503-9222, Fax:81-45-503-9216) CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site for further details. URL:http://genome.gsc.riken.go.jp/ URL:http://fantom.gsc.riken.go.jp/.	COMMENT	
FEATURES		
source		
1. 3287		
/organism="Mus musculus"		
/mol_type="mRNA"		
/strain="C57BL/6J"		
/db_xref="PANTOM_DB:C130052021"		
/db_xref="taxon:10090"		
/clone="C130052021"		
/tissue_type="head"		
/clone_lib="RIKEN full-length enriched mouse cDNA library"		
/dev_stage="16 days embryo"		
622..2871		
misc_feature		
622..2871		
/note="SEMAPHORIN 3B PRECURSOR (SEMAPHORIN A) (SEMA A) (SWISSPROT Q62177, evidence: FASTV, 98.3%ID, 100%length, match=4247) putative"		
3268..3273		
polyA_signal		
3268..3273		
/note="putative"		
3287		
polyA_site		
3287		
/note="putative"		
BASE COUNT	678 a 936 c 970 g 703 t	
ORIGIN		
Query Match	22.8%; Score 617; DB 11; Length 3287;	
Best Local Similarity	59.4%; Pred. No. 1.6e-145;	
Matches 1128; Conservative	0; Mismatches 750; Indels 21; Gaps 4;	
QY	359 AGCTCCAGTTATCATACCTCTCTTTTGATGAGGACGGAGTAGCTGTATGTTGGAGCA	418
Db	775 ACCTGCTGTTATGAGCCCTTCTGCTGATGAGGAGCTGACCCCTGTTTGGGGTCT	834
QY	419 AAGGATCATATATTTTCATTCGACCTGGTGAATAT---CAAGGATTTTCAAAAGATTGTG	475
Db	835 GAGAACCAAGTGGCTTCCTCAGCTGGACAACATCAGCAAGCGCAAGAGCTGCC	894
QY	476 TGGCCAGTATCTTACACAGAGAGATGAATGCAAGTGGCTGAAAGAGCATCTCGAAA	535
Db	895 TGGCCCGCCCGCTGGGAATGCGGTGAAGAATGCAACTGGGCGAGGAAGGACATTTGGTACC	954
QY	536 GAATGTGCTAATTTTCATCAAGTACTTTAAGGCATATATCAGACTCTACTTTGTAGCCCTGT	595
Db	955 GAGTCATGAATCTGCGAAGCTCTGCAACACCTACCAACACACCCACTTTGTGGCCCTGT	1014
QY	596 GGAACGGGGGCTTTTCATCCAAATTTGCACCTTACATTTGAAATTTGACATCATCTCTGAGGAC	655
Db	1015 GGCACAGGGGCTTTCCACCCAACTGTGCTCTTGTGGAGGTGGGCCACCGCTGGAGGAA	1074
QY	656 AATATTTTAACTGCGAGAACTCATATTTGAAAAACGGCCGTGGGAAGAGTCCATATGAC	715
Db	1075 CCCATGCTTCAACTGGACCGGAGGAAACTTTGAGGACGGCAAGGGAAGAGCTCTCTTATGAC	1134
QY	716 CTTAAGCTGCTGACAGCATCCCTTTTAAATAGATGAGCAATATATCTCTGGAATCTGAGCT	775
Db	1135 CCAAGGCATCGGCTGCTCTCGTCTGCTGGGGGAAGAACTGTAATCTGGGGTGACAGCA	1194
QY	776 GATTTTATGGGGCGAGAGCTTTGCTTATCTTCGAACTCTTGGGACACCAACCAATCAGG	835
Db	1195 GACCTTATGGGCGGGACTTTTACCATCTTTTGAAGCCTTGTGAGATCCGAGTCTCCGA	1254
QY	836 ACAGAGCAGCATGATTCAGGTGCTCAATGATCCAAAGTTTCATTTAGTGCCTCCACCTCATC	895

Db 1255 ACAGAGCCCCATGATCCCGCTGGCTCAATGAACCAAGTTGTCAAGGTCTTTGGATC 1314
Qy 896 TCAGAGAGTGACAATCTCTCAAGATGACAAAGTATATCTTTCTCGTGAATCAATA 955
Db 1315 CCAGAGAGTGAGAACCTGATGACGATAAATCTATTTCTTCCGCGAGTCCGCTGG 1374
Qy 956 GATGAGAACACTC---TGAAAAAGCTACTCACGCTAGAAATAGGTGACATATGCAAGAT 1012
Db 1375 GAAGCAGCACCAAGCAATGGCGCATGCTGTGTCTGTGTGGCCAGATCTGCAAGAT 1434
Qy 1013 GACTTTGAGGGCAGCAGAAAGTCTGGTGAATAAATGGAACAATCTCTCAAGCTCGTGTG 1072
Db 1435 GACCTGGGTGGCGAGCGAGCTTGGTCAACAAATGGACCAATTTCTGAAGCGGGCTT 1494
Qy 1073 ATTTGCTCAGTCCAGGTCCTCAATGCAATGCACTGACATCTATTTGATGAACTGCAAGATGA 1132
Db 1495 GTGTGCTCAGTACCTGG---AGTTGAGGGTGACACCCACTTTGACCAACTTCAGGATGT 1551
Qy 1133 TTCTTAATGAATCTTTAAAGATCCTTAAAGATCCAGTTGTATATGAGAGTGTTCACGACTTCC 1192
Db 1552 TTCTTCTGTCTCCGAGAGCCGACAGACACTCTTCTATGCTGTCTTCCACCTCC 1611
Qy 1193 AGTAACATTTTCAAGGGATCAGCGTGTGTATGATAGATGAGTGTGAGAAAGGTG 1252
Db 1612 AGTGTGTCTTCCAGGGCTCTCTGTGTGTGTGTACAGCATGAACGATGTGCGCGAGCC 1671
Qy 1253 TTCTTTGGTCCATATGCCACAGGATGACCCCACTATCAATGATGGTGCCTTATCAAGA 1312
Db 1672 TTCTTTGGGACCTTTTGTCTCACAAAGAGGGGCTTACACACAGTGGGTGTCTTACCAGGGT 1731
Qy 1313 AGAGTCCCTATCCAGCGCCAGGAATTTGTCCAGCAAAACATTTGGTGGTTTGAATCT 1372
Db 1732 CGTGTCCCTACCCAGAGCTGCGATGTGCCCCAGACACTTTGGCACTTCAGCTCC 1791
Qy 1373 ACAAGGACCTTCTGTATGATGTTATACCTTTTGAAGAAGTCAATGAGCCATGATCAAT 1432
Db 1792 ACCAAGGACTTCCAGATGACCTTATCCAGTTTGTGCGAACCACCTCTCATGTACAA 1851
Qy 1433 CCAGTGTTCCTATCAACAATCGCCCAATAGTATCAAAAGGATGTAATATCAATTT 1492
Db 1852 CAGTGTCTGCCATGCGGGGGCGCCCTCTTCTTCAAGTGGAGCTGGGTACACCTTC 1911
Qy 1493 ACACAAATTTGCTAGACCGAGTGTGATGAGAGATGACAGATGATGATGTTATGTTATC 1552
Db 1912 ACCAAATCGCGCAGACCGAGTAGCAGTGCAGTGCAGTACAGATGTTCTCTTCA 1971
Qy 1553 GGAACAGATGTTGGACCGCTTCTTAAAGTAGTTTCAATCTTAAAGGAGACTGTTGATGAT 1612
Db 1972 GGTACAGATGTGGGCACAGTGTGAAAGTGTATCTCAGTCCCCCAAGGCGAGCCTAAT 2031
Qy 1613 TTAGAAGAGTGTCTGCGAAGAAATGACAGTTTTCGGGAACGACTGCTATTTCAGCA 1672
Db 2032 TCTGAAGGACTTCTCTGGAAGAGCTGCGAGGTGTTCGAGGACTCTGCGCTATCACAGC 2091
Qy 1673 ATGAGACTTTTCCACTAAGCAGCAACAATATATATTTGTTTCAACGGCTGGGGTGGCCAG 1732
Db 2092 ATGCAATCTCTCTTAAAGGCAACAATCTACATAGCATCGCGAGCGAGTGGCCAG 2151
Qy 1733 CTCCCTTTACCGGTGTGATATTTACGGGAAGCGTGTGTGAGTGTGCTCGCCCGA 1792
Db 2152 ATTTGCTTGTGATCGTGTGACTGCCCTAGGCGCGCTGCGAGATGCTGTGTCGCCGT 2211
Qy 1793 GACCTTACTGTCTGTTGGATGTTCTCATGTTCTCATGTTTCCGACCTGCAAGAGA 1852
Db 2212 GATCTTACTAGCTTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2271
Qy 1853 CGCAACAGCAGCAGATATAGAAATGGAGACCCACTGACTCTGTTTCAAGCTTACAC 1912
Db 2272 CGTTCCGAGGCAAGATAAGGAATGCGACCCCGACCCCTATGCTC----- 2321
Qy 1913 CATGATATACCATGGCCACAGCCCTGAAGAGAGATCATCTATGTTGTGAGATAGT 1972
Db 2322 ---TGAGACTCTTCTCACTCTGTGCTGTGAGAGAGAGGTTGTTGGGTGTGAGAGCGC 2379

Qy 1973 AGCAATTTTGGATGAGTCCGAAGTCCGAGAGCGCTGCTTATTTGGCAATTCAG 2032
Db 2380 AGCGGTTTCTGGAGTGTGAGCCCGCTGCTCCAGGGCGCATGTGCACTGGACCTTCCA 2439
Qy 2033 AGGCAATTAAGAGAGCGAAAGAGAGATCAGAGTGGATGATCATATCATCAGACAGAT 2092
Db 2440 GGTGAGGGAGGAGGAGCTCACACCAGGTGCTGGCTGAGGAGAGATGAGCGCACTGGC 2499
Qy 2093 CAAGCCCTTCTGCTAGTGTCTACACAGAAAGATTCAGGCAATTAATCTCTGCTCCATGG 2152
Db 2500 CGGGGCTGCTGTTGCGGGGCTGCGCGCGCAGGACTCTGCGGTGTATCTTTTGGTGGCG 2559
Qy 2153 GTGGAACATGTTTCAATAAATCTTCTTAAGGTAAACCTTGAAGTCAATTTGACACAGAG 2212
Db 2560 GTTGAACAAGACTTTTCAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2619
Qy 2213 CATTTGGAAGAACTTCTTCTATAAAGATGATGATGAGAT 2251
Db 2620 CAGGCTGAACGACTGGCAGCGGAGAGGAGGAGCGGCT 2658

RESULT 11
CD349263
LOCUS
DEFINITION
IMAGE: 6854395 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CD349263 729 bp mRNA linear EST 29-MAY-2003
UI-M-FV0-cfg-a-18-0-UI_r1 NIH_BMAP_FV0 Mus musculus cDNA clone
IMAGE: 6854395 5', mRNA sequence.
CD349263
CD349263.1 GI:31140778
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (Bases 1 to 729)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Straubeberg, Ph.D.
Email: rsb@bbs-research.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mousefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pyX-5.
Location/Qualifiers
1..729
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6854395"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_FV0"
/note="Organ: Brain; Vector: pyX-Asc; Site 1: Ecor I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with Ecor I adaptor, digested with NotI and then cloned
directionally into pyX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is AGCGAGAG. This library was created for the University
of Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
Developing Mouse Nervous System', supported by National

FEATURES
source

Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."
BASE COUNT 210 a 165 c 160 g 194 t

Query Match 21.9%; Score 594.6; DB 14; Length 729;
Best Local Similarity 88.5%; Pred. No. 4.4e-140;
Matches 645; Conservative 0; Mismatches 84; Indels 0; Gaps 0;

Qy 1032 GTCTGGTGAATTAATGGAACACATTCCTCAAAGCTCGCTGATTTGCTCAGTGCAGGTC 1091
Db 1 GTCITGTGAATTAATGGAACACATTCCTCAAAGCAGGCTGATTTGCTCAGTGCAGGTC 60

Qy 1092 CAATGGCAATGACACTCATTTTGAATGAATGAGGATGATTTCTAATGAACATTTAAAG 1151
Db 61 CCAATGGCAATGACACCATTTTGAATGAATGAGGATGATTTCTAATGAATCTAAAG 120

Qy 1152 ATCTTAAATCCAGTTGATATGAGGATGATTTACGACTTCAGTAAACATTTTCAAGGAT 1211
Db 121 ATCTTAAATCCAGTTGATATGAGGATGATTTACGACTTCAGTAAACATTTTAAAGGAT 180

Qy 1212 CAGCGTGTGATGATATGAGGATGATTTGAGAGGATGATTTCTGCTCCATATGCC 1271
Db 181 CTGCTGTGATGATATGAGGATGATTTGAGAGGATGATTTCTGCTCCATATGCC 240

Qy 1272 ACAGGATGAGCCCAACTATCAATGGGTGCCTTTATCAAGGAAGATGCTCCCTATCCAGGC 1331
Db 241 ACAGGATGAGCCCAACTATCAATGGGTGCCTTTATCAAGGAAGATGCTCCCTATCCAGGC 300

Qy 1332 CAGGAATGTCAGCAAAACATTTGGTGTGATGATTTGATCTCAAGGACCTTCTGATG 1391
Db 301 CAGGAATGTCAGCAAAACATTTGGTGTGATGATTTGATCTCAAGGACCTTCTGATG 360

Qy 1392 ATGTTAATACCTTTGCAAGAGTCAATCCAGCATGTACAATCCAGTGTTCCTATGAAC 1451
Db 361 ATGTTAATACCTTTGCAAGAGTCAATCCAGCATGTACAATCCAGTGTTCCTATGAAC 420

Qy 1452 ATCCGCCAATAGTATCAAAACGAGTATAATTTATCAAAATTTGCTGAGACC 1511
Db 421 ATCCGCCAATAGTATCAAAACGAGTATAATTTATCAAAATTTGCTGAGACC 480

Qy 1512 GAGTGGATGAGAGATGAGATGATGATTTATCGAAGAGATGTTGGGACG 1571
Db 481 GAGTGGATGAGAGATGAGATGATGATTTATCGAAGAGATGTTGGGACG 540

Qy 1572 TTCTTAAAGTGTTCAGTCCCAAGGAGATGATGATTTGATGATTTGATGATTTGATG 1631
Db 541 TTCTTAAAGTGTTCAGTCCCAAGGAGATGATGATTTGATGATTTGATGATTTGATG 600

Qy 1632 AAGAAATGACAGTATTTTCGGGAACCGATGCTATTTTCAAGCAATGAGGCTTTCCACTAAGC 1691
Db 601 AAGAAATGACAGTATTTTCGGGAACCGATGCTATTTTCAAGCAATGAGGCTTTCCACTAAGC 660

Qy 1692 AGCAACAACTATATTTGGTTCACAGGCTGGGTTGCCAGCTCCCTTTACCGGTTG 1751
Db 661 AGCAACAACTATATTTGGTTCACAGGCTGGGTTGCCAGCTCCCTTTACCGGTTG 720

Qy 1752 ATATTACG 1760
Db 721 ACACTATG 729

RESULT 12
CB196368
LOCUS CB196368 900 bp mRNA linear EST 05-FEB-2003
DEFINITION AGNCOURT_11259678 NIH_MGC_135 Mus musculus cDNA clone.
IMAGE:30138167 5', mRNA sequence.
ACCESSION CB196368
VERSION CB196368.1 GI:28224005
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 900)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDAM0040 row: i column: 24
High quality sequence stop: 655.

FEATURES
source

1..900
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clones="IMAGE:30138167"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_135"
/note="vector: pCMVSPORT6.1; Site 1: EcoRV; Site 2: NotI;
Normalized full-length enriched library from pooled mouse
embryonic limb, maxilla and mandible, day 12.5, 13.5, 14.5
, and 15.5 (size selected for the 0.5-1 kb fragments)
Cloned directionally, priming method: Oligo-dT. cDNA
enrichment: 5x bp, Average insert size 1.6 kb bp.
Normalization (Cot value): 7.5 kb. Priming sequence:
5'-GACTAGTTCATGATCGAGCGCGCCCTTT-3' Tissue contributed by
David Rowe. Library constructed by ResGen, Invitrogen
Corp."

BASE COUNT 286 a 208 c 225 g 178 t 3 others
ORIGIN

Query Match 21.8%; Score 590.4; DB 14; Length 900;
Best Local Similarity 83.0%; Pred. No. 5.6e-139;
Matches 747; Conservative 0; Mismatches 138; Indels 15; Gaps 6;

Qy 1700 CTATATATGTTTCAACGGCTGGGTTGCCAGCTCCCTTTACACCGGTGATATTTAC 1759
Db 1 CTGTACATTTGCTCAACTGCGGAGTGGCAAGCTTCTCTACACCGGTGATATTTAC 60

Qy 1760 GGGAAAGCGTGTGCTGATGTTGCTCGCCGAGACCTTACTGCTGGTGGATGGTCT 1819
Db 61 GGCAAGCGTGTGAGAATGCTGCTCGTGGGACCTTACTGCTGGGATGGGTTCC 120

Qy 1820 GCATGTTCTCGCTATTTTCCACCTGCAAGAGAGCGCACAGCAAGATATAAGAAAT 1879
Db 121 TCATGCTCACGCTATTTTCTTACTGCAAGAGGCGCACAGCAAGATATAAGAAAT 180

Qy 1880 GGAGACCACTGATCTCATGTTTCAGACTT---ACACCATGATATCATCATGGCCACAGC 1936
Db 181 GGAGACCACTGATCTCATGTTTCAGACTTGGAGTGGAGTGGAGTGGAGTGGAGTGG 240

Qy 1937 CCTGAAGAGAGATCATCTATGTTAGAGATAGTAGCATTTTGGATGAGTCCG 1996
Db 241 CTTGAAGAGAGATCATCTATGTTAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 300

Qy 1997 AAGTCGCGAGAGCGCTGGTCTATTGGCAATTCAGAGGGGAAATGAAGCGGAAAAAGAA 2056
Db 301 AAGTCAGAGAGCGCTGGTCTATTGGCAATTCAGAGGGGAAATGAAGCGGAAAAAGAG 360

Qy 2057 GAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAGTCTA 2116
Db 361 GAGATCAGAGTGGTGTATCATATCATCAGGACAGAAACAAAGGCTCTGCTCCGTAGCCTG 420

Qy 2117 CAACAGAGAGTTCAGGCAATTTACCTCTGCCATCGGTTGGACATGGTTCATACAACT 2176
Db 421 CAGAAAGAGGATTCAGGCAATTTACCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480

2177 CTTCTTAAGGTAAACCTGGAAGTCAATTGACACAGAGCAATTTGGAGAACTTCTTCATATA 2236
| | | | |
Db CTTCTTAAGGTAAACCTGGAAGTCAATTGACACAGAGCAATTTGGAGAACTTCTTCATATA 540
| | | | |
2237 GATGATGATGAGATGGCTCTAAGACCAAGAAATGTCAATAGATGACACCTAGCCAG 2296
| | | | |
Db GATGATGATGAGATGGCTCTAAGACCAAGAAATGTCAATAGATGACACCTAGCCAG 600
| | | | |
2297 AAGGTCTGTGTACAGAGATTTCAATGAGCTCATCAACCCCAATCTCAACACAGATGGAT 2356
| | | | |
Db AAGGTCTGTGTACAGAGATTTCAATGAGCTCATCAACCCCAATCTCAACACAGATGGAT 660
| | | | |
2357 GAGTCTGTGTAAACAAGTTTGGAAAAGGACCGAAAACAACGTCGCGCAAGGCGCAGACAT 2416
| | | | |
Db GAGTCTGTGTAAACAAGTTTGGAAAAGGACCGAAAACAACGTCGCGCAAGGCGCAGAC 719
| | | | |
2417 ACCCAGGGAACAGTAACAATGGAAGACCTTACAGAAATAAGAAAGGTAGAAAC-AG 2475
| | | | |
Db ACCCAGGGAACAGTAACAATGGAAGACCTTACAGAAATAAGAAAGGTAGAAAC-AG 779
| | | | |
2476 GAGGACCCACGAATTT--GAGAGGGCACCAGGAGTGTCTGA---GCTGCAATTACCTCT 2529
| | | | |
Db GAGGACCCACGAATTT--GAGAGGGCACCAGGAGTGTCTGA---GCTGCAATTACCTCT 839
| | | | |
2530 AGAAACCTCAAAACAAGTAGAACTTGCCTAGACAATAA----CTGGAATAACAATGCAA 2585
| | | | |
Db AGAAACCTCAAAACAAGTAGAACTTGCCTAGACAATAA----CTGGAATAACAATGCAA 899
| | | | |

RESULT 13

BF667677 889 bp mRNA linear EST 21-DEC-2000
LOCUS 602121773F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4278810 5',
DEFINITION mRNA sequence.
ACCESSION BF667677.1 GI:11941572
VERSION BF667677.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC <http://mgi.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLCW1103 row: a column: 19
High quality sequence stop: 536.
Location/Qualifiers
1. 889
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4278810"
/tissue_type="primitive neuroectoderm"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_56"
/note="Organ: brain; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggccgctggcc); Site 2: SfiI (ggccattatggcc);
Double-stranded cDNA was prepared from cell line RNA.
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGCGCATATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCCGAGGCGGCGCATG-dr(30)BN-3'
(where B = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies

FEATURES
source

1. 889
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4278810"
/tissue_type="primitive neuroectoderm"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_56"
/note="Organ: brain; Vector: pDNR-LIB (Clontech); Site 1:
SfiI (ggccgctggcc); Site 2: SfiI (ggccattatggcc);
Double-stranded cDNA was prepared from cell line RNA.
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CAGCGCATATGGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCCGAGGCGGCGCATG-dr(30)BN-3'
(where B = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies

contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA)."
BASE COUNT 263 a 158 c 250 g 217 t 1 others
ORIGIN
Query Match 21.7%; Score 588.6; DB 10; Length 889;
Best Local Similarity 95.9%; Pred. No. 1.6e-138;
Matches 647; Conservative 0; Mismatches 24; Indels 4; Gaps 4;
QY 1595 AAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAGAAATGACAGTTTTTCGGAA 1654
| | | | |
Db 1 AAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAGAAATGACAGTTTTTCGGAA 60
| | | | |
QY 1655 CCGAGCTGCTATTTTCAGCAATGGAGCTTTCCTACTAAGCAGCAACTATATATGTTTCA 1714
| | | | |
Db 61 CCGAGCTGCTATTTTCAGCAATGGAGCTTTCCTACTAAGCAGCAACTATATATGTTTCA 119
| | | | |
QY 1715 ACGGCTGGGTTGCCAGCTCCCTTTACCGGTTGTATTTTACGGGAAAGCGTGTCT 1774
| | | | |
Db 120 ACGGCTGGGTTGCCAGCTCCCTTTACCGGTTGTATTTTACGGGAAAGCGTGTCT 179
| | | | |
QY 1775 GAGTGTGCTCGCGCGAGACCTTACTGTCTTGGGATGGTTCTGCAATGTTCTGCTAT 1834
| | | | |
Db 180 GAGTGTGCTCGCGCGAGACCTTACTGTCTTGGGATGGTTCTGCAATGTTCTGCTAT 239
| | | | |
QY 1835 TTTCCCACTGCAAGAGACGCAAGAGCAAGCAAGATATAAGAAATGGAGACCCACTGACT 1894
| | | | |
Db 240 TTTCCCACTGCAAGAGACGCAAGAGCAAGCAAGATATAAGAAATGGAGACCCACTGACT 299
| | | | |
QY 1895 CACTGTTTCAGACTTACCAATGATTAATCAACATGCGCCACAGCCCTGGAAGAGAGATCATC 1954
| | | | |
Db 300 CACTGTTTCAGACTTACCAATGATTAATCAACATGCGCCACAGCCCTGGAAGAGAGATCATC 359
| | | | |
QY 1955 TATGCTAGAGATAGTAGCACAATTTTGGAAATGCGAGTCCGAGTCCAGAGACCGTG 2014
| | | | |
Db 360 TATGCTAGAGATAGTAGCACAATTTTGGAAATGCGAGTCCGAGTCCAGAGACCGTG 419
| | | | |
QY 2015 GTCTATTGGCAATTCAGAGCGGAAATGAAGAGCGAAAGAGATGAGAGTGGATGAT 2074
| | | | |
Db 420 GTCTATTGGCAATTCAGAGCGGAAATGAAGAGCGAAAGAGATGAGAGTGGATGAT 479
| | | | |
QY 2075 CATATCATCAGGACAGATCAAGGCGCTTCTGCTAGTGTCTTACCAAGAGAGATTCAGGC 2134
| | | | |
Db 480 CATATCATCAGGACAGATCAAGGCGCTTCTGCTAGTGTCTTACCAAGAGAGATTCAGGC 538
| | | | |
QY 2135 AATTACCTCTGCGATGGGTGGAGATGGTTTCATACAACTCTTCTTAAGTAAACCTG 2194
| | | | |
Db 539 AATTACCTCTGCGATGGGTGGAGATGGTTTCATACAACTCTTCTTAAGTAAACCTG 597
| | | | |
QY 2195 GAAGTCATTGACACAGAGCATTTTGGAAAGAACTTCTTCAAGAGATGATGAGGATGGC 2254
| | | | |
Db 598 GAAGTC-CTGGCAGAGAGCTTTGGGAAAGAAATTTCTTCTTAAGATGATGAGGATGG 656
| | | | |
QY 2255 TCTAAGACCAAGAA 2269
| | | | |
Db 657 TCTTAAAGACCAAGAA 671
| | | | |

RESULT 14

BU703431
LOCUS BU703431
DEFINITION UI-M-FOO-bzo-k-13-0-UI.r1 NIH BMAP_F00 Mus musculus cDNA clone
IMAGE: 6405252 5', mRNA sequence.
ACCESSION BU703431
VERSION BU703431.1 GI:23629234
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 775)
AUTHORS NIH-MGC <http://mgi.nci.nih.gov/>.

with EcoR I adaptor, digested with NotI and then cloned directionally into pVX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is ACGGAGACAG. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 197 a 165 c 156 g 180 t 1 others
ORIGIN

Query Match 20.9%; Score 567.4; DB 14; Length 699;
Best Local Similarity 88.3%; Pred. No. 3.6e-133;
Matches 616; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY	605	GCTTTTCATCCAAATGGACCTACATTTGAATGGACATCATCTCGGAGCAATATTTT	664
Db	1	GCTTTTCATCCAAATGGACCTATATTGAATGGACATCATCTCGGAGCAATATTTT	60
QY	665	AAGCTGGAGAACTCACTATTTTGAACCGCGTGGAGAGTCCATATGACCCCTAAGCTG	724
Db	61	AAGCTGGAGAACTCACTATTTTGAACCGTCTGGAGAGCCCTTATGATCCCAACTA	120
QY	725	CTGACAGCATCCCTTTTAAATAGATGGAGAAATATATCTCTGGAACCTGCAGCTGATTTATG	784
Db	121	CTGACTGCTCTCTTCTAATAGACGGTGAGTTGTACTCTGGAACCTGCTGGGACTTCATG	180
QY	785	GGGCGAGACTTTGCTATCTTCGAACTCTTGGGCAACACCCCAATCAGGACAGAGCAG	844
Db	181	GGACGGGACTTTCGCTATCTTCAGAACACTGGGGGCAACCATCACCCCATCAGGACGGAGCAG	240
QY	845	CATGATCCAGGTGCTCAATGATCCAAAGTTTCAATAGTGGCCACCTCATCTCAGAGAGT	904
Db	241	CATGACTCCCGTGGCTCAATGATCCTAGATTCATGAGTGGCCATCTCATCCAGAGAGT	300
QY	905	GACAACTCTGAGAGTCAAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGGAGAA	964
Db	301	GACAACTCTGAGAGTCAAAAGTATATTTTCTCCGAGAAATGCAATAGACGAGAA	360
QY	965	CACTCTGGAAGAAGCTACTCACGCTAGATAGTAGTCAAGATGCAAGAATGACTTTGGAGGG	1024
Db	361	CACTCTGGAAGAAGCCACTCATGCTAGAATAGGTCAATATGCAAGAATGACTTTGGTGA	420
QY	1025	CACAGAAGTCTGGTCAATAAATGGACAACTTCCTCAAAGCTCGTCTGATTTGCTCAGTG	1084
Db	421	CACAGAAGTCTTGTGAATAATGGACAACTTCCTCAAAGCCGCTGATTTGCTCTGTG	480
QY	1085	CCAGGTCCAAATGGCAATGACACTCATTTTGTATGAATCGAGGATGATTCCTAATGAAC	1144
Db	481	CCCGTCCCAATGGCAATGACACCCATTTTGAATGAATTCAGGATGATTCCTAATGAAC	540
QY	1145	TTTAAAGATCCTAAAATCCAGTTGATATGGAGTGTATGACCTTCAGTAACTATTTTC	1204
Db	541	TCTAAAGATCCTAAAATCCAGTCTATGGAGTGTTCACAACATCAAGCAACATCTTT	600
QY	1205	AAGGATCAGCGCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1264
Db	601	AAGGATCTGCTGT	660
QY	1265	TATGCCACAGGATGGACCCCACTATCAATGGGTGCC	1302
Db	661	TATGCTCACAGATGGTCCCACTATCAGTGGGTGCC	698

Search completed: July 31, 2003, 17:17:23
Job time : 5037 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 11:35:10 ; Search time 679 Seconds
(without alignments)
10769.923 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709

Sequence: 1 aatctttatttattcgtg.....aggcttttttcttaataacc 2709

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq 19Jun03.*

```

1: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*
25: /SIDS1/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2709	100.0	2709	22 AAH47049	Semaphorin D cDNA
2	2508.4	92.6	2530	19 AAX75767	Human semaphorin I
3	2508.4	92.6	2530	24 ABS76512	cDNA encoding huma
4	2508.4	92.6	2601	16 AAO87442	Human semaphorin I
5	1415.2	52.2	1481	16 AAO92331	Human collapsin cD
6	718.6	26.5	2331	20 AAX89112	Human brain tissue
7	718.6	26.5	3871	21 AAA37109	Human PRO1491 (UNQ
8	718.6	26.5	3871	22 AAS46098	Human DNA encoding

9	718.6	26.5	3871	22 AAF54421	DNA encoding prote
10	718.6	26.5	3871	25 ACA57856	Human PRO1491 cDNA
11	718.6	26.5	3871	25 ABX98326	Human cDNA encodin
12	718.6	26.5	3871	25 ABX98828	Novel human secret
13	718.6	26.5	3871	25 ACA05873	Human secreted/tra
14	718.6	26.5	3871	25 ABX97917	Human PRO polynucl
15	718.6	26.5	3871	25 ABX78701	Human PRO polynucl
16	718.6	26.5	3871	25 ABX75714	Human cDNA encodin
17	718.6	26.5	3871	25 ABX76919	Human PRO polynucl
18	718.6	26.5	3871	25 ABX16759	Human cDNA encodin
19	718.6	26.5	3880	20 AAX89113	Human brain tissue
20	647.4	23.9	4474	25 ACC50287	Breast cancer asso
21	599.4	22.1	2337	22 AAF90251	Nucleotide sequenc
22	599.4	22.1	2337	24 ABA97363	Degenerate nucleot
23	596.4	22.0	2898	19 AAV35367	Human semaphorin e
24	594.8	22.0	4460	20 AAZ28469	Mouse semaphorin H
25	556	20.5	5231	23 ABV23116	Human prostate exp
26	556	20.5	5231	23 ABV28955	Human prostate exp
27	554.4	20.5	2975	20 AAX03792	Human semaphorin E
28	554.4	20.5	2975	21 AAC66800	Clone BR533.4 codi
29	553.6	20.4	3988	20 AAZ28470	Mouse semaphorin H
30	552.8	20.4	5177	24 ABV94769	Human pancreatic c
31	474.8	17.5	2349	22 AAH47791	Novel human protei
32	474.8	17.5	2628	22 AAH47790	Novel human protei
33	473.6	17.5	3568	22 AAH47792	Novel human protei
34	473.6	17.5	4859	21 AAA93109	Human secreted pro
35	460.2	17.0	2340	22 AAF90250	Nucleotide sequenc
36	460.2	17.0	2340	24 ABA97362	Human ZSMF-16 enco
37	454.4	16.8	456	23 ABS45213	Human liver single
38	454.4	16.8	456	24 ABS19795	Human genome-deriv
39	385.8	14.2	1444	21 AAC69627	Human secreted pro
40	374	13.8	3054	24 ABK36165	cDNA sequence #556
41	353	13.0	354	22 AA129182	Colon tumour relat
42	353	13.0	354	25 ABZ33368	Human colon tumour
43	335.4	12.4	496	23 ABS32213	Human liver single
44	335.4	12.4	496	24 ABS07290	Human genome-deriv
45	225.6	8.3	561	24 ABL87727	Human ovarian canc

ALIGNMENTS

RESULT 1

AAH47049
ID AAH47049 standard; DNA; 2709 BP.

XX AC AAH47049;

XX DT 29-OCT-2001 (first entry)

XX DE Semaphorin D cDNA sequence.

XX KW Drug resistance; resistance gene; semaphorin D; B94; mel-14 antigen;
XX KW 24p3; proliferin; maspin; cancer; cytostatic; gene therapy; ss.

XX OS Homo sapiens.

XX PN WO200155455-A2.

XX PD 02-AUG-2001.

XX PF 31-JAN-2001; 2001WO-US03161.

XX PR 31-JAN-2000; 2000US-0179191.

XX PA (MILL-) MILLENNIUM PHARM INC.
XX PA (JINS/) JIN S.

XX PI Jin S;

XX XX WPI; 2001-488799/53.

XX PT Determining if a compound modulates the drug resistance of a cell,

PT comprises determining the expression or activity level of a resistance
XX sequence in a cell in the presence of the test compound -

PS Example 1; Fig 1A-B; 79pp; English.

XX The invention relates to a method of determining whether a test compound
CC modulates the drug resistance of a cell that comprises determining the
CC expression or activity level of resistance genes (e.g. semaphorin D, B94,
CC mel-14 antigen, 24p3, proliferin or maspin) in a cell in the presence of
CC the test compound, and comparing its expression or activity level in a
CC cell without the test compound. The drug resistance sequences are useful
CC in identifying drug resistant cells, in screening methods directed to the
CC identification of compounds that can modulate the drug resistance of a
CC cell type or multiple cell types. An isolated resistance protein can be
CC used as an immunogen to generate antibodies that bind the resistance
CC protein. Resistance nucleic acids may be inserted into vectors and used
CC as gene therapy vectors. An anti-resistance protein antibody may be used
CC to isolate a resistance protein, or facilitate the purification of
CC natural resistance protein from cells and of recombinantly produced
CC resistance protein expressed in host cells. The methods are useful for
CC treating a subject having a disorder, such as a drug-resistance cancer,
CC characterized by aberrant resistance sequence expression or activity by
CC administering to the subject a resistance modulator. The present sequence
CC represents a semaphorin cDNA sequence, whose expression was increased in
CC drug resistant EMT6 tumours.

XX Sequence 2709 BP; 831 A; 560 C; 620 G; 698 T; 0 other;

SQ Query Match 100.0%; Score 2709; DB 22; Length 2709;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AATCTTTTATTCGATGTTAAACGCTTAACTAGTATCGATCCCGTGGGGTGTGCG	60
DB	1	AATCTTTTATTCGATGTTAAACGCTTAACTAGTATCGATCCCGTGGGGTGTGCG	60
QY	61	ACCCAGCGTCCGGGAGTAGTGTAGCTCGCTGTTCTCCCATTTGTCAGCCAGTCTATT	120
DB	61	ACCCAGCGTCCGGGAGTAGTGTAGCTCGCTGTTCTCCCATTTGTCAGCCAGTCTATT	120
QY	121	CCAGATTTGTTGAATCTCTCGCCGCAACAATACAGGAAGGAACTAAAGCAGCAAAAG	180
DB	121	CCAGATTTGTTGAATCTCTCGCCGCAACAATACAGGAAGGAACTAAAGCAGCAAAAG	180
QY	181	GACCTACAGCGTCTGCAGCATGGCTGGTTAACTAGGATTTCTGTTTCTGGGGAGT	240
DB	181	GACCTACAGCGTCTGCAGCATGGCTGGTTAACTAGGATTTCTGTTTCTGGGGAGT	240
QY	241	ATTACTTACAGCAAGAGCAAACTATCAGAAATGGGAAGAACTATGCCAAGGCTGAAAT	300
DB	241	ATTACTTACAGCAAGAGCAAACTATCAGAAATGGGAAGAACTATGCCAAGGCTGAAAT	300
QY	301	ATCCTACAAAGAAATGTTGGAAATCCAAATGTGATCACTTTCAATGGCTTGGCCACAG	360
DB	301	ATCCTACAAAGAAATGTTGGAAATCCAAATGTGATCACTTTCAATGGCTTGGCCACAG	360
QY	361	CTCCAGTTATCATCTCTCTTGGATCAGNAACGAGTAGCTGTATGTTGGAGCAA	420
DB	361	CTCCAGTTATCATCTCTCTTGGATCAGNAACGAGTAGCTGTATGTTGGAGCAA	420
QY	421	GGATCACATATTTTCAATCGACTGGTTAAATCAAGGATTTTCAAAAGATTTGTGGCC	480
DB	421	GGATCACATATTTTCAATCGACTGGTTAAATCAAGGATTTTCAAAAGATTTGTGGCC	480
QY	481	AGTATCTTACACAGAAGAGATGAATGCAAGTGGGCTGGAAAGACATCTCTGAAAGAATG	540
DB	481	AGTATCTTACACAGAAGAGATGAATGCAAGTGGGCTGGAAAGACATCTCTGAAAGAATG	540
QY	541	TGCTAAATTTTCATCAGGTAATTAAGGCATATTAATCAGACTCACTTGTAGCCCTGGAC	600
DB	541	TGCTAAATTTTCATCAGGTAATTAAGGCATATTAATCAGACTCACTTGTAGCCCTGGAC	600
QY	601	GGGGGCTTTTTCATCCAAATTTGCACCTACATTTGAAATTTGACATCATCTCTGAGCAATAT	660

1681	TTCCA	CTAAGCAGCAACAACTATAAT	TGGTTCAACGGCTGGGGTGTGCCAGCTCCCTTT	1741
Qy	1741	ACACCGGTGTGATATTTACGGGAAGCGTGTGCTAGTGTGGCTCGCCGAGACCCCTTA	1800	
Db	1741	ACACCGGTGTGATATTTACGGGAAGCGTGTGCTAGTGTGGCTCGCCGAGACCCCTTA	1800	
Qy	1801	CTGTGCTTGGATGGTTCTGCATGTCCTCGCTATTTTCCACTTTCCTCAAGAGACGCAACG	1860	
Db	1801	CTGTGCTTGGATGGTTCTGCATGTCCTCGCTATTTTCCACTTTCCTCAAGAGACGCAACG	1860	
Qy	1861	ACGACAGATATTAAGAAATGGAGACCCACCTGACTCTGTTCAGACTTACACCATGATAA	1920	
Db	1861	ACGACAGATATTAAGAAATGGAGACCCACCTGACTCTGTTCAGACTTACACCATGATAA	1920	
Qy	1921	TCACCATGGCCACAGCCCTGAAGAGAAATCATCTATGTGTAGAGAAATAGTAGCACATT	1980	
Db	1921	TCACCATGGCCACAGCCCTGAAGAGAAATCATCTATGTGTAGAGAAATAGTAGCACATT	1980	
Qy	1981	TTTGGAAATGCAGTCCGAAGTCGACAGAGCGCTGGTCTATTGGCAATTCAGAGGGCGAAA	2040	
Db	1981	TTTGGAAATGCAGTCCGAAGTCGACAGAGCGCTGGTCTATTGGCAATTCAGAGGGCGAAA	2040	
Qy	2041	TGAAGCGCAAAAGACAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCCT	2100	
Db	2041	TGAAGCGCAAAAGACAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCCT	2100	
Qy	2101	TCTGCTACGTAGTCTACAAAGAAAGATTCAGGCAATTACTCTGCGCATCGGTGGACA	2160	
Db	2101	TCTGCTACGTAGTCTACAAAGAAAGATTCAGGCAATTACTCTGCGCATCGGTGGACA	2160	
Qy	2161	TGGTTTCATAACAACTCTTCTTAAGGTAAACCTTGGAAAGTCAATTGACACAGAGCATTTGGA	2220	
Db	2161	TGGTTTCATAACAACTCTTCTTAAGGTAAACCTTGGAAAGTCAATTGACACAGAGCATTTGGA	2220	
Qy	2221	AGAACTTTCTATAAAGATGATGAGAGTGGCTCTTAAAGCAAAAGAAATGTCCAATAG	2280	
Db	2221	AGAACTTTCTATAAAGATGATGAGAGTGGCTCTTAAAGCAAAAGAAATGTCCAATAG	2280	
Qy	2281	CATTGACACCTAGCCAGAAGTCTGGTACAGAGACTTCTGAGCTCATCAACACCCCAA	2340	
Db	2281	CATTGACACCTAGCCAGAAGTCTGGTACAGAGACTTCTGAGCTCATCAACACCCCAA	2340	
Qy	2341	TCTCAACACGATGGATGAGTCTGTGAAACAAGTTTGGAAAAGGACCGAATAAAGTCTG	2400	
Db	2341	TCTCAACACGATGGATGAGTCTGTGAAACAAGTTTGGAAAAGGACCGAATAAAGTCTG	2400	
Qy	2401	GCAAGGCCAGGACATACCCAGGGAACAGTAACAATGGAAGCACTTACAGAAATAA	2460	
Db	2401	GCAAGGCCAGGACATACCCAGGGAACAGTAACAATGGAAGCACTTACAGAAATAA	2460	
Qy	2461	GAAGGTAGAAAACAGGAGGACCCACGAATTTGAGAGGCGACCCAGGAGTGTCTGAGCTGC	2520	
Db	2461	GAAGGTAGAAAACAGGAGGACCCACGAATTTGAGAGGCGACCCAGGAGTGTCTGAGCTGC	2520	
Qy	2521	ATTACCTCTAGAAAACCTCAAAACAGTAGAAACTTGGCTAGACAAATACCTGGAAAAACAA	2580	
Db	2521	ATTACCTCTAGAAAACCTCAAAACAGTAGAAACTTGGCTAGACAAATACCTGGAAAAACAA	2580	
Qy	2581	TGCATATACATGAACTTTTTCATGGCATTTATGTGATGTTTACATGTGTGGAAATTC	2640	
Db	2581	TGCATATACATGAACTTTTTCATGGCATTTATGTGATGTTTACATGTGTGGAAATTC	2640	
Qy	2641	AGCTGAGTTCACCAATATATAATTAATCCATGAGTAACCTTCTTAATAGGCTTTTTT	2700	
Db	2641	AGCTGAGTTCACCAATATATAATTAATCCATGAGTAACCTTCTTAATAGGCTTTTTT	2700	
Qy	2701	CCTAATACC	2709	
Db	2701	CCTAATACC	2709	

RESULT 2
AAX75767

AAX75767 standard; DNA; 2530 BP.
 AAX75767;
 22-JUL-1999 (first entry)
 Human semaphorin III DNA.
 Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer; frameshift mutation; age-related disease; neurodegenerative disorder; Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal; Huntington's disease; multiple sclerosis; alcoholic liver disease; diabetes mellitus type II; microtubule associated protein; Tau; Big Tau; ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M; neurofilament-F; presenilin I; presenilin II; cellular tumour antigen; glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1; bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A; high mobility group protein-C; neuroendocrine specific protein A; ss.
 Homo sapiens.
 WO9845322-A2.
 15-OCT-1998.
 02-APR-1998; 98WO-IB00705.
 10-APR-1997; 97US-0043163.
 (UYUT-) RIJXSUNIV UTRECHT.
 (ROYA-) ROYAL NETHERLANDS ACADEMIES OF SCI.
 (UYRO-) UNIV ROTTERDAM ERASMUS.
 Burbach JPH, Grosveld FG, Van Leeuwen FW;
 WPI; 1998-609901/51.
 Diagnosing disease by detecting frameshift mutations in RNA or corresponding protein mutations - used to diagnose cancer and neurological diseases, particularly Alzheimer's disease, and also for treatment and prevention with specific ribozymes or wild-type RNA
 Disclosure; Figure 16; 258pp; English.
 This invention describes a novel method for the diagnosis of a disease caused by, or associated with, an RNA molecule that has a frameshift mutation. The method is used to diagnose age-related diseases, especially cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's disease, Down's syndrome, myotonic dystrophy, Huntington's disease, multiple sclerosis, alcoholic liver disease, diabetes mellitus type II and many others listed) or susceptibility to these disorders. The method allows a definitive diagnosis of Alzheimer's disease in living patients, at an early stage. It is based on the observation that disease may be caused by mutations in RNA rather than DNA. The invention describes the use of neuronal system RNA molecules, specifically proteins including beta-amyloid precursor protein (beta-APP), the microtubule associated proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule associated protein 2 (MAP2), neurofilament-L, neurofilament-M, neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group protein-C (HMGP-C) and neuroendocrine specific protein A. This sequence encodes the wild type and mutant protein fragments represented in AAY21264-Y21348.
 Sequence 2530 BP; 786 A; 518 C; 576 G; 650 T; 0 other;
 Query Match 92.6%; Score 2508.4; DB 19; Length 2530;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2509; Conservative 0; Mismatches 1; Indels. 0; Gaps 0;
 193 CTGACGATGGGCTGGTTAACTAGGATTCCTGCTCTTCTCGGGAGTATTACTTACAGC 252

FT	Key	Location/Qualifiers	
FT	CDS	16..2331	
FT		/tag= a	
FT		/product= human semaphorin III	
XX			
XX	W09507706-A1.		
XX			
XX	23-MAR-1995.		
XX			
XX	13-SEP-1994;	94WO-US10151.	
XX			
XX	13-SEP-1993;	93US-0121713.	
XX			
XX	(REGC) UNIV CALIFORNIA.		
XX			
XX	Bentley DR, Goodman CS, Kolodkin AL, Matthes D;		
XX	O'Connor T;		
XX			
XX	WPI; 1995-131177/17.		
XX	P-PSDB; AAR71380.		
XX			
XX	New class of semaphorin peptide(s) and polypeptide(s) - are		
XX	potent modulators of nerve cell growth and regeneration		
XX			
XX	Example 2; Page 60-63; 101pp; English.		
XX			
XX	The sequence of the cDNA encoding the human semaphorin III protein.		
XX	The proteins encoded by the grasshopper semaphorin I (AA087441), human		
XX	semaphorin III, vaccinia virus semaphorin IV (AA087443), Drosophila		
XX	semaphorin I and II (AA087444-5), Tribolium semaphorin I (AA087446) or		
XX	varicella major (smallpox) virus semaphorin IV (AA087447) genes were used		
XX	to generate a series of peptides (AAR70370-R70418), which retain		
XX	semaphorin receptor binding activity. The semaphorin derived or		
XX	semaphorin receptor derived peptides are potent modulators of nerve cell		
XX	growth, immune responsiveness and viral pathogenesis. They can be used		
XX	in diagnosis and treatment of neurological disease and		
XX	neuro-regeneration, immune modulation and diagnosis and treatment of		
XX	viral and oncological infection and diseases.		
XX	(Updated on 25-MAR-2003 to correct PN field.)		
XX			
XX	Sequence 2601 BP; 809 A; 533 C; 593 G; 666 T; 0 other;		
XX			
XX	Query Match	92.6%; Score 2508.4; DB 16; Length 2601;	
XX	Best Local Similarity	100.0%; Pred. No. 0;	
XX	Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;		
QY	193	CTGCAGCATGGCTGTTAACTAGGATGTCGTCTTTCTGGGAGTATTACTTACAGC	252
DB	9	CTGCAGCATGGCTGTTAACTAGGATGTCGTCTTTCTGGGAGTATTACTTACAGC	68
QY	253	AAGAGCAAACTATCAGAATGGGAAGAACAAATGTGCCAAGGCTGAAATTTATCTACAAAGA	312
DB	69	AAGAGCAAACTATCAGAATGGGAAGAACAAATGTGCCAAGGCTGAAATTTATCTACAAAGA	128
QY	313	AATGTTGAATCCAAATGTGATCACTTTCAATGGCTTGGCCACAGCTCCAGTTATCA	372
DB	129	AATGTTGAATCCAAATGTGATCACTTTCAATGGCTTGGCCACAGCTCCAGTTATCA	188
QY	373	TACCTTCCTTTTGGATGGACGAGTAGGCTGTATGTTGGAGCAAGGATCACAATTT	432
DB	189	TACCTTCCTTTTGGATGGACGAGTAGGCTGTATGTTGGAGCAAGGATCACAATTT	248
QY	433	TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACAC	492
DB	249	TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACAC	308
QY	493	CAGAAGAGATGAATCGAATGGCTGGAAAGACATCTCGAAGAAATGTGCTAATTTCAAT	552
DB	309	CAGAAGAGATGAATCGAATGGCTGGAAAGACATCTCGAAGAAATGTGCTAATTTCAAT	368
QY	553	CAAGGTACTTAAGGCATATATACACTCTTGTACGCTGTGAAACGGGGGCTTTTCA	612
DB	369	CAAGGTACTTAAGGCATATATACACTCTTGTACGCTGTGAAACGGGGGCTTTTCA	428

QY	613	TCCAAATTTGCACCTACATTTGAAATTTGGACATCATCTCAGGACAAATATTTTAACTGGA	672
DB	429	TCCAAATTTGCACCTACATTTGAAATTTGGACATCATCTCAGGACAAATATTTTAACTGGA	488
QY	673	GAACTCACATTTTGAAGACGGCCGCGGGAAGTCCATATGACCTTACCTAGCTCTCAGACG	732
DB	489	GAACTCACATTTTGAAGACGGCCGCGGGAAGTCCATATGACCTTACCTAGCTCTCAGACG	548
QY	733	ATCCCTTTTAATAGATGAGAAATATATCTCTGGAACTCAGCTGATTTTATGGGCGGAGA	792
DB	549	ATCCCTTTTAATAGATGAGAAATATATCTCTGGAACTCAGCTGATTTTATGGGCGGAGA	608
QY	793	CTTTGCTATCTTCGGAATCTTTGGGCACCAACCAATCAGGACAGAGCATGATTC	852
DB	609	CTTTGCTATCTTCGGAATCTTTGGGCACCAACCAATCAGGACAGAGCATGATTC	668
QY	853	CAGTGGCTCAATGATCCAAAGTTTCATTTAGTGGCCACCTCATCTCAGAGAGTGACAATCC	912
DB	669	CAGTGGCTCAATGATCCAAAGTTTCATTTAGTGGCCACCTCATCTCAGAGAGTGACAATCC	728
QY	913	TGAAGATGACAAAGTATATCTTTCTCCGTGAAATGCAATAGATGAGAACACTCTGG	972
DB	729	TGAAGATGACAAAGTATATCTTTCTCCGTGAAATGCAATAGATGAGAACACTCTGG	788
QY	973	AAAAGCTACTCACGCTAGATAGGTTCAGATATGCAAGAAATGACTTTTGAGGGGCACAGAAG	1032
DB	789	AAAAGCTACTCACGCTAGATAGGTTCAGATATGCAAGAAATGACTTTTGAGGGGCACAGAAG	848
QY	1033	TCTGTTGAATTAATGGACAAATCTCTCAAAAGCTCGTCTGATTTGCTCAGTGGCCAGTCC	1092
DB	849	TCTGTTGAATTAATGGACAAATCTCTCAAAAGCTCGTCTGATTTGCTCAGTGGCCAGTCC	908
QY	1093	AAATGGCAATTCACACTCAATTTTGAATGAACTCAGAGTATTTCTTAAATGAACTTTAAAGA	1152
DB	909	AAATGGCAATTCACACTCAATTTTGAATGAACTCAGAGTATTTCTTAAATGAACTTTAAAGA	968
QY	1153	TCCTAAATAATCCAGTTGTATATGGAGTGTTTTACGACTTCCAGTAACATTTTCAAGGGATC	1212
DB	969	TCCTAAATAATCCAGTTGTATATGGAGTGTTTTACGACTTCCAGTAACATTTTCAAGGGATC	1028
QY	1213	AGCGTGTGATGATAGCATGATGATGAGAGGGTGTTCCTTGTTCATATGCCCCA	1272
DB	1029	AGCGTGTGATGATAGCATGATGATGAGAGGGTGTTCCTTGTTCATATGCCCCA	1088
QY	1273	CAGGATGGACCCCAACTATCAATGGTGGCTTTTCAAGGAAGAGTCCCTATCCACGGCC	1332
DB	1089	CAGGATGGACCCCAACTATCAATGGTGGCTTTTCAAGGAAGAGTCCCTATCCACGGCC	1148
QY	1333	AGGAACCTTGTCCAGCAAAACATTTTGGTGGTGTGACTCTCAAAAGGACCTTCTCTGATGA	1392
DB	1149	AGGAACCTTGTCCAGCAAAACATTTTGGTGGTGTGACTCTCAAAAGGACCTTCTCTGATGA	1208
QY	1393	TGTTATAACCTTTGCAAGAGTCAATCCAGCCATGTACAATCCAGTGTTCCTATGAACAA	1452
DB	1209	TGTTATAACCTTTGCAAGAGTCAATCCAGCCATGTACAATCCAGTGTTCCTATGAACAA	1268
QY	1453	TCGCCCAATAGTATCAAAACGGATGTAATTTATCAATTTTACAAATTTGCTAGACCG	1512
DB	1269	TCGCCCAATAGTATCAAAACGGATGTAATTTATCAATTTTACAAATTTGCTAGACCG	1328
QY	1513	AGTGGATGCAAGATGGAAGTATGATGTTATGTTTATCGGAAACAGATGTTGGACCGT	1572
DB	1329	AGTGGATGCAAGATGGAAGTATGATGTTTATGTTTATCGGAAACAGATGTTGGACCGT	1388
QY	1573	TCCTTAAAGTAGTTTCAATTCCTAAGGAGACTGTTGATGATTTAGAGAGGTTCTGCTGGA	1632
DB	1389	TCCTTAAAGTAGTTTCAATTCCTAAGGAGACTGTTGATGATTTAGAGAGGTTCTGCTGGA	1448
QY	1633	AGAAATGACAGTGTTCGGGAAACCGACTGTCTATTTTACGCAATGAGCTTTCCTAAGCA	1692
DB	1449	AGAAATGACAGTGTTCGGGAAACCGACTGTCTATTTTACGCAATGAGCTTTCCTAAGCA	1508

```
QY 1693 GCACAACTATATATTTGGTTCAACGGCTGGGGTTGCCAGCTCCCTTTACACGGGTGCA 1752
Db 1509 GCAACAACTATATATTTGGTTCAACGGCTGGGGTTGCCAGCTCCCTTTACACGGGTGCA 1568
QY 1753 TATTTACGGGAAGCGTGTGCTGAGTGTGGTTCCTCGCCCGAGACCTTACTGTGCTGGGA 1812
Db 1569 TATTTACGGGAAGCGTGTGCTGAGTGTGGTTCCTCGCCCGAGACCTTACTGTGCTGGGA 1628
QY 1813 TGGTTCCTGCATGTTCTCGCTATTTTCCCACTGCTGCAAGAGACGCACAGACGACAGATAT 1872
Db 1629 TGGTTCCTGCATGTTCTCGCTATTTTCCCACTGCTGCAAGAGACGCACAGACGACAGATAT 1688
QY 1873 AAGAAATGGAGACCCACTGACTCACTGTGTTACAGACTTACACCATGATAATCACCATGGCCA 1932
Db 1689 AAGAAATGGAGACCCACTGACTCACTGTGTTACAGACTTACACCATGATAATCACCATGGCCA 1748
QY 1933 CAGCCCTGAAGAGAGAAATCATCTATGTTGTTAGAGAAATAGTAGCAATTTTGGAAATGCAG 1992
Db 1749 CAGCCCTGAAGAGAGAAATCATCTATGTTGTTAGAGAAATAGTAGCAATTTTGGAAATGCAG 1808
QY 1993 TCCGAAGTCGCAGAGAGCGCTGCTTATTTGGCAATTTCCAGAGCGGAAATGAAGAGCGAAA 2052
Db 1809 TCCGAAGTCGCAGAGAGCGCTGCTTATTTGGCAATTTCCAGAGCGGAAATGAAGAGCGAAA 1868
QY 2053 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAG 2112
Db 1869 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAG 1928
QY 2113 TCTAACAGAGAGATTACAGGCAATTAATCTCTGCGCATGCGGTGGAAATGCGGTTCATACA 2172
Db 1929 TCTAACAGAGAGATTACAGGCAATTAATCTCTGCGCATGCGGTGGAAATGCGGTTCATACA 1988
QY 2173 AACTCTTCTTAAGTAACTGGAAGTCAATTCAGACAGAGCATTTGGAGAACTTCTTCA 2232
Db 1989 AACTCTTCTTAAGTAACTGGAAGTCAATTCAGACAGAGCATTTGGAGAACTTCTTCA 2048
QY 2233 TAAAGATGATGAGAGATGGCTTAAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2292
Db 2049 TAAAGATGATGAGAGATGGCTTAAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2108
QY 2293 CAGAAAGTCTGTACAGAGACTTCAATGAGCTCATCAACACCCCAATCTCAACACGAT 2352
Db 2109 CAGAAAGTCTGTACAGAGACTTCAATGAGCTCATCAACACCCCAATCTCAACACGAT 2168
QY 2353 GATGAGTCTGTGCAACAGTTTGGAAAGGACCGAAACAACTCGCGCAAGGCCAGG 2412
Db 2169 GATGAGTCTGTGCAACAGTTTGGAAAGGACCGAAACAACTCGCGCAAGGCCAGG 2228
QY 2413 ACATACCCAGGGAACAGTAACAAATGGAAGCACTTACAGAAAAATAAGAAAGGTAGAAA 2472
Db 2229 ACATACCCAGGGAACAGTAACAAATGGAAGCACTTACAGAAAAATAAGAAAGGTAGAAA 2288
QY 2473 CAGGAGGCCACGAAATTTGAGAGGCCACCGAGGAGTGTCTGAGCTGCATTAACCTCTAGA 2532
Db 2289 CAGGAGGCCACGAAATTTGAGAGGCCACCGAGGAGTGTCTGAGCTGCATTAACCTCTAGA 2348
QY 2533 AACCTCAAAAGTAGAAGTGTGCTGTAGCAATAAATGGAAGAAACAAATGCAATATACAT 2592
Db 2349 AACCTCAAAAGTAGAAGTGTGCTGTAGCAATAAATGGAAGAAACAAATGCAATATACAT 2408
QY 2593 GAACTTTTTTCATGGCATTTATGTTGATTTTCAATTTGGTGGAAATTCAGCTGAGTTCCA 2652
Db 2409 GAACTTTTTTCATGGCATTTATGTTGATTTTCAATTTGGTGGAAATTCAGCTGAGTTCCA 2468
QY 2653 CCAATTTAAATTAATCCATGAGTAACTTTTCCCTAATAGGCTTTTTTCC 2702
Db 2469 CCAATTTAAATTAATCCATGAGTAACTTTTCCCTAATAGGCTTTTTTTC 2518
```

RESULT 5

AAQ92331.

ID AAQ92331 standard; cDNA; 1481 BP.

XX

```
AC AAQ92331;
XX 01-NOV-1995 (first entry)
XX Human collapsin cDNA.
XX Collapsin; antibody; therapy; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 50..1480
XX /*tag= a
XX US5416197-A.
XX 16-MAY-1995.
XX 15-OCT-1993; 93US-0136922.
XX 15-OCT-1993; 93US-0136922.
XX (UYPE-) UNIV PENNSYLVANIA.
XX Luo Y, Raper JA;
XX WPI; 1995-193478/25.
XX P-PSDB; AAR74175.
XX
XX New antibody to human collapsin - used to inhibit the activity of
XX collapsin, to induce neurite out-growth and to treat individuals with
XX nerve damage.
XX Disclosure; Columns 11-16; 11pp; English.
XX
XX Human collapsin and its encoding nucleic acid may be used to
XX identify agents which modulate the ability of human collapsin to
XX collapse the growth cone of neurons. An antibody capable of
XX specifically binding at least a portion of the collapsin protein can
XX be used to purify human collapsin and to inhibit the activity of the
XX protein. It can be used to induce neurite outgrowth by neuronal
XX cells and to treat individuals suffering from nerve damage.
XX
XX Sequence 1481 BP; 454 A; 299 C; 345 G; 383 T; 0 other;
XX
Query Match 52.2%; Score 1415.2; DB 16; Length 1481;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1420; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 899 GAGAGTGACATCTCTGAAGATGACAAAGTATACATTTTCTTCCTGAAAAATGCAATAGAT 958
Db 53 GAACATGACATCTCTGAAGATGACAAAGTATACATTTTCTTCCTGAAAAATGCAATAGAT 112
QY 959 GGAGAACACATCTGAAAAAGCTACTCAGCTAGAAATAGGTACAGATATGCAAGAAATGACTTT 1018
Db 113 GGAGAACACATCTGAAAAAGCTACTCAGCTAGAAATAGGTACAGATATGCAAGAAATGACTTT 172
QY 1019 GGAGGGACAGAACTCTGGTGAATAAATGGACACATCTCTCAAGCTCGTCTGATTTGC 1078
Db 173 GGAGGGACAGAACTCTGGTGAATAAATGGACACATCTCTCAAGCTCGTCTGATTTGC 232
QY 1079 TCAGTGGCAGGTCCAAATGGCATTGACACTCATTTTGTATGAACCTGCAGAGATGATTCCTA 1138
Db 233 TCAGTGGCAGGTCCAAATGGCATTGACACTCATTTTGTATGAACCTGCAGAGATGATTCCTA 292
QY 1139 ATGAACCTTTAAAGATCCTAAAAATCCAGTGTGTATATGAGAGTGTATGAGCTTCCAGTAAC 1198
Db 293 ATGAACCTTTAAAGATCCTAAAAATCCAGTGTGTATATGAGAGTGTATGAGCTTCCAGTAAC 352
QY 1199 ATTTTCAAGGATCAGCGGTGTGTATATGAGATGATGTGATGATGATGATGATGATGATGAT 1258
Db 353 ATTTTCAAGGATCAGCGGTGTGTATATGAGATGATGTGATGATGATGATGATGATGATGATGAT 412
```


QY 626 TACATTGAAATGGACATCATCTGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTT 685
Db 469 TATATTGATCTTGGAGTCTCAAGAGGAGATATTTATTTCAAACTAGACACAGTAAATTG 528
QY 686 GAAACGGCGGTGGGAAGAGTCCATATGACCCCTAAGCTGTCGACAGCATCCCTTTTAAATA 745
Db 529 GAGCTGGCAGAGCTGAAATGTCTTTGATCTCTCAGCAGCCCTTTGCTTCAGTAATGACA 588
QY 746 GATGGAGAAATTAATCTCTGGAATCGCAGCTGATTTTATGGGGGAGAGCTTTGCTATCTTC 805
Db 589 GATGAGTACCTCTACTCTGGAACAGCTTCTGATTTCTTTGGCAAGATATCTGCAATCACT 648
QY 806 CGAATCTTTGGGC-----ACCACCCCAATCAGGACAGCAGCAGCATGATCCAGG 856
Db 649 CGATCCCTTTGGGCTTACTCATGACCCACCTACTCATCAGAACTGACATTTTCAGAGCACTAC 708
QY 857 TGGCTCAATGATCCAAAGTTTCATTTAGTGGCCACCTCATCTCAGAGAGTGCACATCTCGAA 916
Db 709 TGGCTCAATGGAGCAAAATTTATTTGAACTTTCTTCATACCAGACACCTTACAAATCCAGAT 768
QY 917 GATGACAAAGTATATCTTTTCTTCGTGAATATGCAATAGATGGAGAACACTCTGGAAAA 976
Db 769 GATGATAAAATATATTTCTTTCTTCTGTAATCATCTCAAGAAAGGCAAGTACCTCCGATAAA 828
QY 977 GCTACTCAGCTAGATAGTGCAGATATGCAAGATGATTTGGAGGGCAGAGATCTG 1036
Db 829 ACCATCTTTCTGAGTGGAGAGTGTGAAGATGATGAGGAGGACAAGCGAGCTG 888
QY 1037 GTGAATAAATGGACAACTTCCTCAAGGCTCGTCTGATTTGCTCAGTGGCAGGTCCTCAAT 1096
Db 889 ATAAACAAGTGGAGCACTTTCTTAAGGCCAGACTGATTTGCTCAATTTCTGGAAGTGAT 948
QY 1097 GGCATTCACACTCATTTTGAAGAACTGCGAGATGATTTCTTAATGAACCTTTAAAGATCT 1156
Db 949 GGGGAGATATCTTCTGATGAGCTTCAAGATATTTATTTACTCCCCCAAGAGATGAA 1008
QY 1157 AAAAATCCAGTTGATATGAGTGTGTTACGACTTCAGGTAACAATTTTCAAGGATCAGCC 1216
Db 1009 AGAAATCTGTAGTATGAGTCTTCTTACCAACGAGTCTTCTTCAAGGCTCTGCT 1068
QY 1217 GTGTGATGATAGCATGATGATGAGAAGGGTGTCTTGGTCCATATGCCCCACAGG 1276
Db 1069 GTTTGTGTATAGCATGGCTGACATCAGAGCAGTGTAAATGCTCCATATGCTCATAAG 1128
QY 1277 GATGGACCCAACTATCAATGGGTGCTTATCAGGAAAGTCCCTATCCAGGGCCAGGA 1336
Db 1129 GAAAGTGCAGACCATCGTTGGGTGCAGTATGATGGGAGAAATTCCTTATCCAGGCTGT 1188
QY 1337 ACTGTGCCAGCAAAACATTTG--GTGGTGTGACTCTACAAAGGACCTTCTCTGATGAT 1393
Db 1189 ACATGTCCCAAGCAAAACCTATGACCACTGATTAAGTCCACCGAGATTTCCAGATGAT 1248
QY 1394 GTTATAACCTTTGCAAGAAAGTCTATCCAGCCATGTACAATCCAGTGTTCCTATGAACAAT 1453
Db 1249 GTCATCAGTTTCATAAAGGGCACTCTGTGATGTATAAGTCCGTATACCCAGTTGCAGGA 1308
QY 1454 CGCCCAATAGTATCAAAACGGATGTAATATTATCAATTTACAAATTTGCTAGACCGA 1513
Db 1309 GGACCAACGTTCAAGAGAAATCAATGTGGATTACAGACTGCACACAGATAGTGGTGATCAT 1368
QY 1514 GTGGATGCAGAGATGCAGATGATGATTTATGTTTATCGGAACAGATGTTGGAGCCGTT 1573
Db 1369 GTCAATTCGAGAGATGGCCAGTACGATGTAATGTTTCTTGGAAACAGCAATGGAACTGTC 1428
QY 1574 CTTAAAGTATGTTCAATTTCTTAAGGAGACTTGGTATGATTTAGAAGAGGTTCTGCTGGAA 1633
Db 1429 CTCAAAGTTGTGAGCATTTCAAAGGAAAGTGG--AATATGGAAGAGGTAGTGTGGAG 1485
QY 1634 GAAATGACATGTTTTCGGGNAACGACTGCTATTTTACGCAATGGAGCTTTCCACTAAGCAG 1693
Db 1486 GAGTTGCAGATATTTCAAGCACTCATCAATCATCTTGAAATGGAATTTGCTCTGAAGCAG 1545

QY 1694 CAACAACTATATATTGGTTTCAACGGCTGGGGTTCGCCAGCTCCCTTTTACACCGGTGTGAT 1753
Db 1546 CAACAACTTGTACATTTGGTTCCCGAGATGATTTAGTTAGCTCTCTCTGACAGATGCGAC 1605
QY 1754 ATTTACGGAAAGCGTGTGCTGAGTGTTCCTCGCCGAGAGCCCTTACTGTGCTTGGGAT 1813
Db 1606 ACTTATGGAAAGCTTGGCAGACTGTTGTCTTGGCAGAGACCCCTACTGTGCTCTGGGAT 1665
QY 1814 GGTCTGCTGATGTTCTCGCTATTTTCCCACTGCAAGAGACGACACAAGACGAAGATATA 1873
Db 1666 GGAATGATGCTCTCGATATGCTCTACTTCTTAAAGAGAGCTAGAGCCCAAGATGTA 1725
QY 1874 AGAATGAGAGACCCACTGACTCAGTCTGTTTACAGACTTACACCATGATAATCACCATGGCCAC 1933
Db 1726 AAATATGGGACCCCAATCACCAGTGTGGGACATCGAAGACAGCATTAGTCAATG--AA 1782
QY 1934 AGCCCTGAAGAGAGAAATCATCTATGCTGTAGAGAAATAGTAGACATTTTTTGAATGCACT 1993
Db 1783 ACTGCTGATGAAAGGTTGATTTTGGCATTTGAATTTAACTCAACCTTTTCTGGAATGTATA 1842
QY 1994 CGAAGTCCGAGAGAGCGCTGCTATTTGGCAATTTCCAGAGCGGAAATGAAGACGAAAA 2053
Db 1843 CCTAAATCCCAACAGCAACTATTTAAATGGTATATCCAGAGGTCAAGGGATGAGCATCGA 1902
QY 2054 GAAGAGATCAGATGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGCTAGT 2113
Db 1903 GAGGATTTGAAGCCGATGAAGAAATCATCAAAAGGAAATATGGGCTACTGATTCGAGT 1962
QY 2114 CTACAACAGAAGGATTCAGGCAATTAATCTGCTGCAATGCGGTGGAAACATGGGTTTCAACAA 2173
Db 1963 TTGCAAGAAGAAGGATTTCTGGGATGATTTACTGCAAAAGCCAGGAGCACACTTTTCATCCAC 2022
QY 2174 ACTCTTCTTAAGGTAACCTCGAAGTCAATTCACACAGAGCATTTTGAAGAATCTTCTTCAT 2233
Db 2023 ACCATAGTGAAGTCACTTTGAATGTCAATGAGAAATGAACAGATGGAATAATCACCAGAGG 2082
QY 2234 AAAGATGATGATG 2246
Db 2083 GCAGAGCATGAGG 2095
RESULT 7
AAA37109
ID AAA37109 standard; cDNA; 3871 BP.
XX
AC AAA37109;
XX
DT 08-AUG-2000 (first entry)
XX
Human PRO1491 (UNQ760) cDNA sequence SEQ ID NO:309.
DE Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;
KW transmembrane; secretion; immunoadhesion; pharmaceutical; screening;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200012708-A2.
XX
PD 09-MAR-2000.
XX
PF 01-SEP-1999; 99WO-US20111.
XX
PR 01-SEP-1998; 98US-0098716.
PR 01-SEP-1998; 98US-0098749.
PR 01-SEP-1998; 98US-0098750.
PR 02-SEP-1998; 98US-0098803.
PR 02-SEP-1998; 98US-0098821.
PR 02-SEP-1998; 98US-0098843.
PR 09-SEP-1998; 98US-0098536.
PR 09-SEP-1998; 98US-0098596.
PR 09-SEP-1998; 98US-0098598.
PR 09-SEP-1998; 98US-0098602.

```
PR 09-SEP-1998; 98US-0099642.
PR 10-SEP-1998; 98US-0099741.
PR 10-SEP-1998; 98US-0099754.
PR 10-SEP-1998; 98US-0099763.
PR 10-SEP-1998; 98US-0099792.
PR 10-SEP-1998; 98US-0099808.
PR 10-SEP-1998; 98US-0099812.
PR 10-SEP-1998; 98US-0099815.
PR 10-SEP-1998; 98US-0099816.
PR 15-SEP-1998; 98US-0100385.
PR 15-SEP-1998; 98US-0100388.
PR 15-SEP-1998; 98US-0100390.
PR 16-SEP-1998; 98US-0100584.
PR 16-SEP-1998; 98US-0100627.
PR 16-SEP-1998; 98US-0100661.
PR 16-SEP-1998; 98US-0100662.
PR 16-SEP-1998; 98US-0100664.
PR 17-SEP-1998; 98US-0100683.
PR 17-SEP-1998; 98US-0100684.
PR 17-SEP-1998; 98US-0100710.
PR 17-SEP-1998; 98US-0100711.
PR 17-SEP-1998; 98US-0100919.
PR 17-SEP-1998; 98US-0100930.
PR 18-SEP-1998; 98US-0100848.
PR 18-SEP-1998; 98US-0100849.
PR 18-SEP-1998; 98US-0101014.
PR 18-SEP-1998; 98US-0101068.
PR 18-SEP-1998; 98US-0101071.
PR 22-SEP-1998; 98US-0101279.
PR 23-SEP-1998; 98US-0101471.
PR 23-SEP-1998; 98US-0101472.
PR 23-SEP-1998; 98US-0101474.
PR 23-SEP-1998; 98US-0101475.
PR 23-SEP-1998; 98US-0101476.
PR 23-SEP-1998; 98US-0101477.
PR 23-SEP-1998; 98US-0101479.
PR 24-SEP-1998; 98US-0101738.
PR 24-SEP-1998; 98US-0101741.
PR 24-SEP-1998; 98US-0101743.
PR 24-SEP-1998; 98US-0101915.
PR 24-SEP-1998; 98US-0101916.
PR 29-SEP-1998; 98US-0102207.
PR 29-SEP-1998; 98US-0102240.
PR 29-SEP-1998; 98US-0102307.
PR 29-SEP-1998; 98US-0102330.
PR 29-SEP-1998; 98US-0102331.
PR 30-SEP-1998; 98US-0102484.
PR 30-SEP-1998; 98US-0102487.
PR 30-SEP-1998; 98US-0102570.
PR 30-SEP-1998; 98US-0102571.
PR 01-OCT-1998; 98US-0102684.
PR 01-OCT-1998; 98US-0102687.
PR 02-OCT-1998; 98US-0102965.
PR 06-OCT-1998; 98US-0103258.
PR 06-OCT-1998; 98US-0103449.
PR 07-OCT-1998; 98US-0103314.
PR 07-OCT-1998; 98US-0103315.
PR 07-OCT-1998; 98US-0103328.
PR 07-OCT-1998; 98US-0103395.
PR 07-OCT-1998; 98US-0103396.
PR 07-OCT-1998; 98US-0103401.
PR 08-OCT-1998; 98US-0103673.
PR 08-OCT-1998; 98US-0103678.
PR 08-OCT-1998; 98US-0103679.
PR 08-OCT-1998; 98US-0103711.
PR 14-OCT-1998; 98US-0104257.
PR 20-OCT-1998; 98US-0104987.
PR 20-OCT-1998; 98US-0105000.
PR 20-OCT-1998; 98US-0105002.
PR 21-OCT-1998; 98US-0105104.
PR 22-OCT-1998; 98US-0105169.
PR 22-OCT-1998; 98US-0105266.
PR 26-OCT-1998; 98US-0105693.

PR 26-OCT-1998; 98US-0105694.
PR 27-OCT-1998; 98US-0105807.
PR 27-OCT-1998; 98US-0105881.
PR 27-OCT-1998; 98US-0105882.
PR 27-OCT-1998; 98US-0106062.
PR 28-OCT-1998; 98US-0106023.
PR 28-OCT-1998; 98US-0106029.
PR 28-OCT-1998; 98US-0106030.
PR 28-OCT-1998; 98US-0106032.
PR 28-OCT-1998; 98US-0106033.
PR 29-OCT-1998; 98US-0106178.
PR 29-OCT-1998; 98US-0106248.
PR 29-OCT-1998; 98US-0106384.
PR 29-OCT-1998; 98US-0108500.
PR 30-OCT-1998; 98US-0106464.
PR 03-NOV-1998; 98US-0106856.
PR 03-NOV-1998; 98US-0106902.
PR 03-NOV-1998; 98US-0106905.
PR 03-NOV-1998; 98US-0106919.
PR 03-NOV-1998; 98US-0106932.
PR 03-NOV-1998; 98US-0106934.
PR 10-NOV-1998; 98US-0107783.
PR 17-NOV-1998; 98US-0108775.
PR 17-NOV-1998; 98US-0108779.
PR 17-NOV-1998; 98US-0108787.
PR 17-NOV-1998; 98US-0108788.
PR 17-NOV-1998; 98US-0108801.
PR 17-NOV-1998; 98US-0108802.
PR 17-NOV-1998; 98US-0108806.
PR 17-NOV-1998; 98US-0108807.
PR 17-NOV-1998; 98US-0108867.
PR 17-NOV-1998; 98US-0108925.
PR 18-NOV-1998; 98US-0108848.
PR 18-NOV-1998; 98US-0108849.
PR 18-NOV-1998; 98US-0108850.
PR 18-NOV-1998; 98US-0108851.
PR 18-NOV-1998; 98US-0108852.
PR 18-NOV-1998; 98US-0108858.
PR 18-NOV-1998; 98US-0108904.
XX
XX (GETH ) GENENTECH INC.
XX
XX Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;
XX WPI; 2000-237871/20.
XX P-PSDB; AAY99427.
XX
XX New mammalian DNA sequences encoding transmembrane, receptor or
XX secreted PRO polypeptides, useful for screening of potential peptide or
XX small molecule inhibitors of the relevant receptor/ligand interactions
XX
XX Claim 2; Fig 175; 773pp; English.
XX
XX AAA37022 to AAA37144 encode the new isolated human transmembrane,
XX receptor or secreted PRO polypeptides given in AAY99340 to AAY99462. The
XX transmembrane and receptor PRO proteins can be used for screening of
XX potential peptide or small molecule inhibitors of the relevant
XX receptor/ligand interactions. The polypeptides and nucleotide sequences
XX encoding them have various industrial applications, including uses as
XX pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent
XX PCR primers and hybridisation probes used in the isolation of the PRO
XX polypeptides from the present invention.
XX
XX Sequence 3871 BP; 1251 A; 704 C; 805 G; 1111 T; 0 other;

Query Match 26.5%; Score 718.6; DB 21; Length 3871;
Best Local Similarity 61.9%; Pred. No. 4.2e-206;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;

QY 269 AATGGAGACAAATGTGCCAGGCTGAAATATCTTACAAAGAAATCTTGAATCCAC 328
DB 215 ACTTTGAAGCAAAATATTTCCAAGACTCAAGCTACCAAGACTTCTGCTTTCAAT 274
```


KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder;
KW PCR primer.
XX
OS Homo sapiens.
XX WO200168848-A2.
PN
XX
PD 20-SEP-2001.
XX
XX 28-FEB-2001; 2001WO-US06520.
XX
XX 01-MAR-2000; 2000WO-US05601.
XX 02-MAR-2000; 2000WO-US05841.
XX 03-MAR-2000; 2000US-187202P.
XX 06-MAR-2000; 2000US-186968P.
XX 14-MAR-2000; 2000US-189320P.
XX 14-MAR-2000; 2000US-189328P.
XX 15-MAR-2000; 2000WO-US06884.
XX 21-MAR-2000; 2000US-190828P.
XX 21-MAR-2000; 2000US-191007P.
XX 21-MAR-2000; 2000US-191048P.
XX 28-MAR-2000; 2000US-191314P.
XX 29-MAR-2000; 2000US-192655P.
XX 29-MAR-2000; 2000US-193032P.
XX 30-MAR-2000; 2000US-193053P.
XX 04-APR-2000; 2000WO-US08439.
XX 04-APR-2000; 2000US-194449P.
XX 11-APR-2000; 2000US-194647P.
XX 11-APR-2000; 2000US-195975P.
XX 11-APR-2000; 2000US-196000P.
XX 11-APR-2000; 2000US-196187P.
XX 11-APR-2000; 2000US-196690P.
XX 11-APR-2000; 2000US-196820P.
XX 18-APR-2000; 2000US-198121P.
XX 18-APR-2000; 2000US-198585P.
XX 25-APR-2000; 2000US-199397P.
XX 25-APR-2000; 2000US-199550P.
XX 25-APR-2000; 2000US-199654P.
XX 03-MAY-2000; 2000US-201516P.
XX 17-MAY-2000; 2000WO-US13705.
XX 22-MAY-2000; 2000WO-US14042.
XX 30-MAY-2000; 2000WO-US14941.
XX 02-JUN-2000; 2000WO-US15264.
XX 05-JUN-2000; 2000US-209832P.
XX 28-JUL-2000; 2000WO-US20710.
XX 22-AUG-2000; 2000US-0644848.
XX 24-AUG-2000; 2000WO-US23328.
XX 08-NOV-2000; 2000WO-US30952.
XX 01-DEC-2000; 2000WO-US32678.
XX 20-DEC-2000; 2000WO-US34956.
XX
XX (GETH) GENENTECH INC.
XX
XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
XX Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
XX WPI; 2001-602746/68.
XX P-PSDB; AAU29197.
XX
XX Novel nucleic acids encoding PRO polypeptides, used to diagnose the
XX presence of tumours, such as prostate and breast tumours, in mammals and
XX to screen for modulators of the compounds -
XX
XX Claim 2; Fig 347; 774pp; English.
XX
XX Sequences AAS45925-AAS46231 represent DNA molecules encoding and PCR
XX primers for PRO polypeptides of the invention. The sequences of the
XX invention can be used to detect the presence of a tumour in a mammal by
XX comparing the level of expression of a PRO polypeptide in a test sample
XX of cells from the animal and a control sample of normal cells, whereby a
XX higher level of expression in the test sample indicates the presence of a
XX tumour in the mammal. Mammals include dogs, cats, cattle, horses, sheep,
XX pigs, goats and rabbits but are preferably human. The polypeptides can be

CC used to stimulate tumour necrosis factor (TNF) alpha release from human
CC blood, when contacted with it. A specific polypeptide can be used to
CC stimulate the proliferation or differentiation of chondrocyte cells. The
CC PRO proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders.
XX
XX Sequence 3871 BP; 1251 A; 704 C; 805 G; 1111 T; 0 other;

Query Match 26.5%; Score 718.6; DB 22; Length 3871;
Best Local Similarity 61.9%; Pred. No. 4.2e-206;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;

QY 269 AATGGGAAGACAATGCGCAAGGCTGAAATATCTCTACAAAGAAATGTTGGAATCCCAAC 328
DB |||||
QY 215 ACTTTGAAGCAAAATATTCGAAGACTCAAGCTAACCTACCAAGACTTCTGCTTCAAT 274
DB |||||
QY 329 AATGTGATCACTTTCAATGGCTTGGCCAAAGCTCCAGTTATCATACCTTCTCTTTGGAT 388
DB |||||
QY 275 AGCTGATTCCTTTTGGGTTTCATCAGAGGACTGGATTTTCAAACCTCTTCTCTTAGAT 334
DB |||||
QY 389 GAGGAACGAGTAGCTGTATGTTGGACCAAGGATCATATATTTTCAATTCGACCTGGTT 448
DB |||||
QY 335 GAGGAAGAGGACGGCTGCTCTTGGGAGCCAAAGACCAATCTTTTCTACTCAGTCTGGTT 394
DB |||||
QY 449 AA---TATCAAGGATTTTCAAAGATTTGTGGCCAGTATCTTACACAGAGAGAGATGA 505
DB |||||
QY 395 GACTTAAACAAAATTTTAAAGAGATTTATTTGGCTGCTGCAAGGAACGGTGGAAATTA 454
DB |||||
QY 506 TGCAAGTGGGCTGGAAAGACATCTGAAAGAAATGTGCTAAATTTTCATCAAGTACTTAAG 565
DB |||||
QY 455 TGTAAATTTAGCTGGGAAGATGCCAATACAGAAATGTCAAATTTTCATCAGAGTACTTCAG 514
DB |||||
QY 566 GCATATATCAGACTCCTGTAGCCCTGTGGACGGGGCTTTTTCATCCAAATTTGCACC 625
DB |||||
QY 515 CCTATACAAAACTCAGATATATGTGTGGAACTGGAGCATTTTCATCCAAATATGTGG 574
DB |||||
QY 626 TACATTGAAATTTGACATCATCTGAGGACAAATATTTTAAAGCTGGGAACTCACAATTTT 685
DB |||||
QY 575 TATATTTGATCTTGGAGTCTACAGAGGATATTTATTTCAAAGTACACACATATTTTG 634
DB |||||
QY 686 GAAACCGCCGCTGGGAAGAGTCCATATGACCCCTAAGCTGTGTCAGAGCATCCCTTTTAATA 745
DB |||||
QY 635 GAGCTGGCAGACTGAAATGTCCTTTTCGATCCTCAGCAGCCTTTTTCCTCAGTAATGACA 694
DB |||||
QY 746 GATGAGAAATTTATCTCTGGAACCTGCGACTGATTTTATGGGGCGGAGCTTTGCTATCTTC 805
DB |||||
QY 695 GATGAGTACCTCTACTCTGGAAACAGCTTCTGATTTCTTGGCAAGAGATCTGCATTCAT 754
DB |||||
QY 806 CGAATCTTGGGC-----ACCACCCCAATCAGGACAGCAGCATGATTCACAGG 856
DB |||||
QY 755 CGATCCCTTGGGCTCTCTCATGACCACTACATCAGAACTGACATTTTCAGAGCACTAC 814
DB |||||
QY 857 TGGCTCAATGATCCAAAGTTTCATTTAGTCCCACTCATCTCAGAGAGTGACAATCCCTGAA 916
DB |||||
QY 815 TGGCTCAATGAGCAAAATTTTGGAACTTTCTTATACACAGACACCTCAATCCAGAT 874
DB |||||
QY 917 GATGACAAAGTATATCTTTTCTCCGTAAGAAATGCAATAGATGGAGAACACTCTGGAAAA 976
DB |||||
QY 875 GATGATAAAATATATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 934
DB |||||
QY 977 GCTACTCAGCTAGATAGTGCATATGCAAGATGACTTTTGGAGGCGCAGAGTCTG 1036
DB |||||
QY 935 ACCATCTTTCTCGAGTTGGAAGAGTTTGTGAAGATGATGTAGGAGGACACGCGACCTG 994
DB |||||
QY 1037 GTGAATAAATGGCAACAATTTCTCAAAGCTCGTCTGATTTTCTCAGTCCAGGTCCTCAAT 1096
DB |||||
QY 995 ATAAACAGTGGACGACTTTCTTAAAGCCAGACTGATTTGCTCAATTTCTCTGGAGTAT 1054
DB |||||
QY 1097 GGCATTGACACTCAATTTGATGAATCTGAGGATGATTTCTTAAAGATCTTTTAAAGATCCT 1156
DB |||||

Db	1055	GGGGCAGACTACTTACTTTGTATGAGCTTCAAGATATTTATTATTACTTCCCCCAAGAGATGAA	11114
Qy	1157	AAAAATCCAGTTGTATATGAGTGTTTTACGACTTCCAGTAACTATTTTCAAGGGATCAGCC	1216
Db	1115	AGAAATCCTGTAGTATATGGAGTCTTTTACTACAACGAGCTCCATCTTCAAGAGCTCTGCT	1174
Qy	1217	GTGTGTATGTATAGCATGAGTGTGTGAGAAAGGGTGTTCCTTGGTCCATATGCCCAAGG	1276
Db	1175	GTTTGTGTATAGCATGGCTGACATCAGAGCAGTGTTTTAAATGGTCCATATGCTCATAG	1234
Qy	1277	GATGGACCCAACTCATATGGGTGCTTATCAAGGAAGAGTCCCTTATCACGGCCAGGA	1336
Db	1235	GAAAGTGCAGACCACTCGTTGGGTGCAGTATGATGGGAGAAATCCCTATCACGGCCCTGGT	1294
Qy	1337	ACTTGTCCAGCAAAACATTTG---GTGGTTTTGCACCTCAAAAGGACCTTCTCTGATGAT	1393
Db	1295	ACATGTCAGCAAAACCTATGACCCCACTGATTAAGTCCACCCGAGATTTTTCAGATGAT	1354
Qy	1394	GTTATAACCTTTGCAAGAAGTCACTCCAGCCATGTACAATCCAGTGTTTTCCATGCAACAAT	1453
Db	1355	GTCACTAGTTTCATAAAGCGGCACCTCTGTGATGTATAAGTCGTATACCCAGTTGCAGGA	1414
Qy	1454	CGCCCAATAGTGATCAAAACGGATGTAAATATCAATTTACAAATTTGCTGTAGACCGA	1513
Db	1415	GGACCAACGTTTCAAGAGAAATCAATGTGGATTTACAGACTGACACAGATAGTGTGGATCAT	1474
Qy	1514	GTGGATCGAGAAGATGACAGTATGATGTTATGTTTATCGGAACAGATGTTGGGACCGTT	1573
Db	1475	GTCAITGCGAAGATGGCCAGTACGATGTAAATGTTTCTTGGAACAGACATTTGGAACGTGC	1534
Qy	1574	CTTAAAGTAGTTTCAAATTTCTTAAGGAGACTTGGTATGATTAATTTAGAAGAGGTTCTGCTGAA	1633
Db	1535	CTCAAAGTTGTACCAITTTCAAAGGAAAAGTGG--AATATGGAAGAGGTAGTCTGGAG	1591
Qy	1634	GAAATGACAGTTTTTCGGGAACCGACTGTCTATTTTCAGCAATGAGAGCTTTCCACTAAGCAG	1693
Db	1592	GAGTTGCGAGATATTTCAAGACACTCATCAATCATCTTTGAACATGGAAATTTGTCTCTGAAGCAG	1651
Qy	1694	CAACAACTATATATTTGTTCAACGGCTGGGGTTGCCAGCTCCCTTTTACACCGGTGTGAT	1753
Db	1652	CACAAATGTACATTTGGTTTCCGAGATGGATTAGTTCAAGTCTCTCTTTGCAAGATGGCAG	1711
Qy	1754	ATTTTACGGGAAAGCGTGTGCTGAGTGTTCCTCGCCCGGAGACCCCTTACTGTCTGGTGGGAT	1813
Db	1712	ACTTATGGGAAAGCTTGCGCAGACTGTTGTCTTGCCAGAGACCCCTACTGTGCTGGAT	1771
Qy	1814	GGTTCTGCATGTTCTCGTATATTTTCCCACTGTGCAAAAGAGACGCAAGACAGCATATATA	1873
Db	1772	GGAAATGCATGCTCTCGATATGCTCTTACTTCTAAAAGGAGAGCTAGACGCCAAGATGTA	1831
Qy	1874	AGAAATGGAGACCCACTGACTCACTGTTTCAGACTTTACACCATGATATACCATGGCCAC	1933
Db	1832	AAATATGGGACCCCAATCACCCAGTGTGGGACATCGAAGACAGCATTAGTCTATG---AA	1888
Qy	1934	AGCCCTCAAGAGAGAAATCATCTATGGTGTAGAGAAATAGTAGCAATTTTTTGGGAATGCAGT	1993
Db	1889	ACTGCTGATGAAAAGGTGATTTTTTGGCATGGAATTTAACTCAACCTTTCTGGAAATGTATA	1948
Qy	1994	CCGAAGTCGAGAGAGCGCTGGTCTATTGGCAATTCAGAGGGCGAAATGAAGAGCGGAAAA	2053
Db	1949	CCTAAATCCCAACAAGCAAATATTAAATGGTATATCCAGAGGTTCAGGGGATGAGCATCGA	2008
Qy	2054	GAAGAGATCAGATGGGATGATCATATCAGGACAGATCAAGGCCTTCTGCTACGTAGT	2113
Db	2009	GAGGAGTTGAAGCCCGCATGAAGAATCATCAAAACGGAAATATGGGCTACTGATTCGAAGT	2068
Qy	2114	CTACAAACAGAGGATTCAGGCAATTACTCTGCCATGCGGTGGAAACATGGGTTTCATACAA	2173
Db	2069	TTGAGAAAGAGGATTTCTGGATGTATATCTGCAAGCCGAGNGCACATTTTCATCCAC	2128
Qy	2174	ACTCTTCTTAAGGTAAACCTCGGAAGTCAATTGACACAGAGCATTTTGGGAAGCACTTCTTCAT	2233
Db	2129	ACCAATGAGCTGACTTTGAAATGTCAATGAGAAATGAAACAGATGAAATACCCAGAGG	2188

Qy	2234	AAAGTATGATG	2246	
Db	2189	GCAGATGAGG	2201	
RESULT 9				
AA	AAAF54421			
ID	AAAF54421	standard; DNA; 3871 BP.		
AC	AAF54421;			
XX	AAF54421;			
DT	02-APR-2001	(first entry)		
XX				
DE	DNA encoding protein of the invention #86			
XX				
XX	Secreted; transmembrane; gene therapy; ss			
KW				
XX	Unidentified.			
OS				
XX	WO200078961-A1.			
FN				
XX	28-DEC-2000.			
XX				
PF	18-FEB-2000; 2000WO-US04342.			
XX				
PR	23-JUN-1999; 99US-0141037.			
PR	20-JUL-1999; 99US-0144758.			
PR	26-JUL-1999; 99US-0145698.			
PR	01-SEP-1999; 99WO-US20111.			
PR	29-OCT-1999; 99US-0162506.			
PR	30-NOV-1999; 99WO-US28313.			
PR	02-DEC-1999; 99WO-US28551.			
PR	16-DEC-1999; 99WO-US30095.			
PR	05-JAN-2000; 2000WO-US00219.			
PR	06-JAN-2000; 2000WO-US00376.			
XX				
PA	(GETH) GENENTECH INC.			
XX				
PI	Baker KP, Botstein D, Desnoyers L, Ecto			
PI	Gao W, Goddard A, Godowski PJ, Grimaldi			
PI	Pan J, Paoni NF, Roy MA, Smith V, Stev			
PI	Watanabe CK, Williams PM, Wood WT;			
XX				
DR	WPI; 2001-071395/08.			
XX				
PT	Secreted and transmembrane proteins and nu			
PT	useful as hybridization probes, in chromos			
-PT	therapy -			
XX				
PS	Claim 2; Fig 171; 787pp; English.			
XX				
CC	The present invention relates to secreted			
CC	These proteins and the DNA encoding them			
CC	probes, in chromosome and gene mapping and			
CC	anti-sense RNA and DNA. They may also be u			
CC	transgenic animals or knockout animals whic			
CC	development and screening of therapeutical			
CC	The nucleic acids may also be used in gene			
XX				
XX				
SQ	Sequence 3871 BP; 1251 A; 704 C; 805 G; 11			
Query Match 26.5%; Score 718.6;				
Best Local Similarity 61.9%; Pred. No. 4.2;				
Matches 1233; Conservative 0; Mismatches				
Qy	269	AATGGGAAGACAATGTGCGCAAGGCTGGAATTTAT		
Db	215	ACTTTGAGCAAAATATCCAGAGCTCAAGCTT		
Qy	329	AATGTGATCACTTTCAATGGCTTGGCCACACAG		
Db	275	AGCTGATATCCCTTTTGGGTTTCATCAGAAGG		

PN US2003036143-A1.
XX 20-FEB-2003.
XX
XX
XX 02-JUL-2002; 2002US-0187600.
XX
XX 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32578.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-US22639.
PR 18-SEP-1997; 97US-US22666.
PR 17-OCT-1997; 97US-US22509.
PR 21-OCT-1997; 97US-US24869.
PR 24-OCT-1997; 97US-US263120P.
PR 24-OCT-1997; 97US-US263121P.
PR 28-OCT-1997; 97US-US263540P.
PR 28-OCT-1997; 97US-US263541P.
PR 28-OCT-1997; 97US-US263544P.
PR 28-OCT-1997; 97US-US263564P.
PR 29-OCT-1997; 97US-US263734P.
PR 31-OCT-1997; 97US-US263870P.
PR 31-OCT-1997; 97US-US264103P.
PR 13-NOV-1997; 97US-US265311P.
PR 21-NOV-1997; 97US-US266120P.
PR 24-NOV-1997; 97US-US266466P.
PR 24-NOV-1997; 97US-US266772P.
PR 11-DEC-1997; 97US-US269335P.
PR 12-DEC-1997; 97US-US269425P.
PR 17-DEC-1997; 97US-US269870P.
PR 18-DEC-1997; 97US-US268017P.
PR 10-MAR-1998; 98US-US27450P.
PR 11-MAR-1998; 98US-US27632P.
PR 11-MAR-1998; 98US-US276499P.
PR 20-MAR-1998; 98US-US278886P.
PR 20-MAR-1998; 98US-US278939P.
PR 27-MAR-1998; 98US-US279664P.
PR 27-MAR-1998; 98US-US279786P.
PR 31-MAR-1998; 98US-US280107P.
PR 31-MAR-1998; 98US-US280194P.
PR 01-APR-1998; 98US-US280327P.
PR 01-APR-1998; 98US-US280333P.
PR 08-APR-1998; 98US-US281049P.
PR 08-APR-1998; 98US-US281070P.
PR 09-APR-1998; 98US-US281195P.
PR 15-APR-1998; 98US-US281838P.
PR 21-APR-1998; 98US-US282568P.
PR 21-APR-1998; 98US-US282569P.
PR 22-APR-1998; 98US-US282704P.
PR 22-APR-1998; 98US-US282797P.
PR 28-APR-1998; 98US-US28322P.
PR 28-APR-1998; 98US-US283495P.
PR 29-APR-1998; 98US-US283496P.
PR 29-APR-1998; 98US-US283499P.
PR 29-APR-1998; 98US-US283559P.
PR 05-MAY-1998; 98US-US284366P.
PR 06-MAY-1998; 98US-US284414P.
PR 07-MAY-1998; 98US-US284639P.
PR 07-MAY-1998; 98US-US284640P.
PR 07-MAY-1998; 98US-US284643P.
PR 15-MAY-1998; 98US-US285579P.
PR 15-MAY-1998; 98US-US285808P.
PR 15-MAY-1998; 98US-US28582P.
PR 15-MAY-1998; 98US-US285700P.
PR 18-MAY-1998; 98US-US286023P.
PR 22-MAY-1998; 98US-US286392P.
PR 22-MAY-1998; 98US-US286486P.
PR 28-MAY-1998; 98US-US287098P.
PR 28-MAY-1998; 98US-US287208P.
PR 02-JUN-1998; 98US-US287609P.
PR 02-JUN-1998; 98US-US287759P.
PR 03-JUN-1998; 98US-US287827P.
PR 04-JUN-1998; 98US-US288025P.
PR 04-JUN-1998; 98US-US288028P.
PR 04-JUN-1998; 98US-US288029P.
PR 04-JUN-1998; 98US-US288033P.
PR 04-JUN-1998; 98US-US288326P.
PR 05-JUN-1998; 98US-US288167P.
PR 05-JUN-1998; 98US-US288202P.
PR 05-JUN-1998; 98US-US288212P.
PR 05-JUN-1998; 98US-US288217P.
PR 09-JUN-1998; 98US-US288655P.
PR 10-JUN-1998; 98US-US288722P.
PR 10-JUN-1998; 98US-US288738P.
PR 10-JUN-1998; 98US-US288740P.
PR 10-JUN-1998; 98US-US288811P.
PR 10-JUN-1998; 98US-US288824P.
PR 10-JUN-1998; 98US-US288825P.
PR 10-JUN-1998; 98US-US288826P.
PR 11-JUN-1998; 98US-US288861P.
PR 11-JUN-1998; 98US-US288863P.
PR 11-JUN-1998; 98US-US288876P.
PR 12-JUN-1998; 98US-US289090P.
PR 12-JUN-1998; 98US-US289105P.
PR 16-JUN-1998; 98US-US289512P.
PR 16-JUN-1998; 98US-US289514P.
PR 17-JUN-1998; 98US-US289538P.
PR 17-JUN-1998; 98US-US289598P.
PR 17-JUN-1998; 98US-US289653P.
PR 18-JUN-1998; 98US-US289908P.
PR 19-JUN-1998; 98US-US289952P.
PR 22-JUN-1998; 98US-US290246P.
PR 22-JUN-1998; 98US-US290252P.
PR 22-JUN-1998; 98US-US290254P.
PR 24-JUN-1998; 98US-US290429P.
PR 24-JUN-1998; 98US-US290435P.
PR 24-JUN-1998; 98US-US290444P.
PR 24-JUN-1998; 98US-US290461P.
PR 24-JUN-1998; 98US-US290535P.
PR 24-JUN-1998; 98US-US290540P.
PR 25-JUN-1998; 98US-US290676P.
PR 25-JUN-1998; 98US-US290678P.
PR 25-JUN-1998; 98US-US290688P.
PR 25-JUN-1998; 98US-US290690P.
PR 25-JUN-1998; 98US-US290694P.

Db 935 ACCATCTTCTCGAGTGGAGAGTTTGTAAAGATGATGTAGGAGCAACAGCGCTG 994
Qy 1037 GTCAATAAATGACACATCTCTCAAGCTCGTCTGATTTGCTCAGTGCAGGTCAAT 1096
Db 995 ATAAACAAGTGGACGACTTTTCTTAAGGCGAGACTGATTTGCTCAATCTCTGGAATGAT 1054
Qy 1097 GGCATTGACACTCATTTTGTATGAATGCGAGGATGATCTCTAATGAACCTTTAAAGATCCT 1156
Db 1055 GGGGAGATACTTACTTTGATGAGCTTCAAGATATTTATTTACTCCCAACAGAGTGA 1114
Qy 1157 AAAATCCAGTTGTATATGAGAGTTTACGACTTCAGTAAACATTTTCAAGGATCAGCC 1216
Db 1115 AGAAATCCTGTAGTATATGAGTCTTTACTTAAACAGCTCCATCTTCAAGGCTCTGCT 1174
Qy 1217 GTGTGTATATAGCATGATGATGTGAGAGGGTGTCTTGGTCCATATGCCACAGG 1276
Db 1175 GTTTGTGTATAGCATGCTGACATCAGACGAGTTTAAATGGTCCATATCTCATAAG 1234
Qy 1277 GATGGACCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTATCCAGGCCAGGA 1336
Db 1235 GAAAGTGCAGACCATGTTGGGTGAGTATGATGGAGATTCCTTATCCAGGCTGT 1294
Qy 1337 ACTTGTCCAGCAAAACATTG---GTGGTTTGACTCTACAAAGAACCTTCTCTGATGAT 1393
Db 1295 ACATGTCCAAAGCAAAACCTATGACCCACTGATTAAGTCCACCGAGATTTTCCAGATGAT 1354
Qy 1394 GTTATACCTTTGCAAGAGTATCCAGCCATGATCAATCCAGTGTCTCTATGAACAT 1453
Db 1355 GTCATGTTCATAAAGCGGCACTCTGTGATGATTAAGTCCGTATACCCCATGTCAGGA 1414
Qy 1454 CGCCCAATAGTATCAAAACGGATGTAATTTATCAATTTACACAAATTCGTGAGACGA 1513
Db 1415 GGACCAACGTTCAAGAGATCAATGTGGATTACAGACTGACACAGATAGTGTGATCAT 1474
Qy 1514 GTGGATGCAAGAGATGGAAGATGATGTTATGTTTATCGGAACAGATTTGGGACCGT 1573
Db 1475 GTCAATGCAAGAGATGGCCAGTACGATGTAATGTTCTTGGAAACAGACATTCGAATGTC 1534
Qy 1574 CTTAAGTATGTTCAATCTTAAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGA 1633
Db 1535 CTCAAAGTTGTGAGCATTTCAAGGAAAGTGG---AATAGGAAGAGGATGTGTGGAG 1591
Qy 1634 GAAATGACAGTGTTCGGGAACCGACTGTCTTTCAGCAATGAGGCTTTCACCTAAGCAG 1693
Db 1592 GAGTTCAGATATTCAAGCACTCATCAATCACTTGAACATGGAATTTCTCTGAGCAG 1651
Qy 1694 CAACAACTATATTTGGTTCAAGGCTGGGGTTGCCAGCTCCCTTTTACCCGGTGTGAT 1753
Db 1652 CAACAACTTGTATTTGGTTCCGAGATGATGATTTAGTTTCACTCTCTTGCACAGATCGGAC 1711
Qy 1754 ATTTACGGGAAGCGTGTCTGATGTTGCTCGCCGAGACCTTACTGTCTTGGAT 1813
Db 1712 ACTTATGGGAAGCTTGGCAGACTGTGTCTTGGCAGAGACCTTACTGTGCTGGGAT 1771
Qy 1814 GGTTCGTGATGTTCTCGCTATTTTCCCACTGCAAGAGACGCAAGACGACAAATATA 1873
Db 1772 GGAATATGATGCTCTCGATATGCTCTACTCTTAAAGGAGAGCTAGAGCCCAAGATGTA 1831
Qy 1874 AGAATGGAGACCACTGATCTCATCTGTTGAGACTTACACCATGATTAATCACCATGGCCAC 1933
Db 1832 AAATATGGGACCAATCAACCCAGTCTGGGACATCGAAGACAGCATAGTATG---AA 1888
Qy 1934 AGCCCTGAAGAGAGATCATCTATGTTGTAGAGATAGTAGCACAATTTTGAATGTCAGT 1993
Db 1889 ACTGCTGATGAAGAGTGAATTTTGGCATATGAATTTAACTCAACCTTCTGGAATGTATA 1948
Qy 1994 CGCAAGTCCAGAGCGCTGTCTTATTTGGCAATTCAGAGGCGAAATGAAGAGCGAA 2053
Db 1949 CCTAAATCCCAACAGCACTATTAATGGTATATCCAGAGTCAAGGATGAGCATCGA 2008
Qy 2054 GAAGAGATCAGAGTGGATGATCATATCATCAGACAGATCAAGGCGCTTCTGCTAAGT 2113

Db 2009 GAGGAGTTGAAGCCCGATGAAGAATCATCAAAACGGAATATGGCTACTGATTCGAAGT 2068
Qy 2114 CTACAACAAGAGGATTCAGGCAATTTACCTCTGCCATGCGGTGGAAACATGGGTTTATACAA 2173
Db 2059 TTGCAAGAAGAGGATTCGGGATGTTTACTTGAAGCCAGGACACACTTTTCAATCCAC 2128
Qy 2174 ACTCTTCTTAAGGTAAACCTCGAAAGTCTATTGACACAGAGCATTTTGGAGAAGACTTTCTTCA 2233
Db 2129 ACCATAGTGAAGTGCACCTTTGATGATGTCATTCAGATGAACAGATGGAATAATACCCAGAGG 2188
Qy 2234 AAAGATGATGATG 2246
Db 2189 GCAGAGCATGAGG 2201

RESULT 12
ABX98828
ID ABX98828 standard; cDNA; 3871 BP.
XX AC ABX98828;
XX 20-MAY-2003 (first entry)
XX DE Novel human secreted and transmembrane protein PRO1491 cDNA.
KW Human; secreted protein; transmembrane protein; cytostatic;
KW gene therapy; TNF-Agonist-Alpha; chondrocyte stimulator; tumour;
KW adrenal tumour; lung tumour; colon tumour; breast tumour;
KW prostate tumour; rectal tumour; cervical tumour; liver tumour;
KW gene; ss.
XX OS Homo sapiens.
XX PN US2003036157-A1.
XX PD 20-FEB-2003.
XX PF 02-JUL-2002; 2002US-0188769.
XX PR 16-SEP-1998; 98WO-US19330.
PR 07-OCT-1998; 98WO-US21141.
PR 01-DEC-1998; 98WO-US25108.
PR 08-MAR-1999; 99WO-US05028.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 01-DEC-1999; 99WO-US28301.
PR 02-DEC-1999; 99WO-US28551.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 15-MAR-2000; 2000WO-US06884.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.

PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063240P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063714P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-064311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066772P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077649P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081338P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083322P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083599P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084539P.
PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 09-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.

PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089553P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090698P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 01-SEP-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 02-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 10-SEP-1998; 98US-099754P.

Query Match

26.5%; Score 718.6; DB 25; Length 3871;

Best Local Similarity 61.9%; Pred. No. 4.2e-206;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;

QY	269	AATGGGAAGACAAATGTCGAGGCTGAATATCTACAAAGAAATGTTGGAAATCCAC	328	Db	1235	GAAAGTCAGACACCATCGTTGGGTGTCAGTATGATGGGAGAAATTCCTTATCCACGGCTGCT	1294
Db	215	ACTTTGAAGCAAAATATCCAGACTCAAGCTAACTCAAGAGACTTGTCTTTCAAAT	274	QY	1337	ACTTGCTCCAGCAAAACATTTG---GTGTTTTGACTCTACAAAGAGACCTTCTCTGATGAT	1393
QY	329	AATGTGATCACATTTCAATGGCTTGGCCAAACAGCTTATCATACCTTCTCTTTGGAT	388	Db	1295	ACATGTCGAAGCAAAACCTATGACCCACTGATTAAGTCCACCCGAGATTTTCCAGATGAT	1354
Db	275	AGCTGTATTCCTTTTGGGTTTCATCAGAAGNCTGGATTTTCAAACTCTTCTTAGAT	334	QY	1394	GTTATAACCTTTTGCAGAAGTCAATCCAGCATGTAAATCCAGTGTTCCTATGAACAAT	1453
QY	389	GAGGAACGAGTAGGCTGTATGTTGGAGCAAGAGATCAATATTTTCAATTCGACCTGGTT	448	Db	1355	GTCAATCAGTTTCATAAAGCGGCACCTCTGTGATGATATAAGTCGTATACCCAGTTGCAGGA	1414
Db	335	GAGGAAGAGGCGAGGCTGCTCTTGGGAGCCAAAGACCATCTTTCTACTCAGTCTGTT	394	QY	1454	CGCCCAATAGTATCAAAACCGGATGTAATTAATCAATTTACAAATTTGTCTGTAGACGA	1513
QY	449	AA---TATCAAGATTTTCAAAAGATTTGTGGCCAGTATCTTACACAGAGAGATGAA	505	Db	1415	GGACCAACGTTTCAAGAGAAATCAATGTGTGATTAACAGACTGACACAGATGTGTGTGATCAT	1474
Db	395	GACTTAACCAAAATTTTAAGAGATTTATTTGGCCTGCTGCAAGGAACGGGTGGAATTA	454	QY	1514	GTGGATGACAGAAGATGGACAGTATGATGTTATTTTTCGGAAACAGATGTTTGGGACCGTT	1573
QY	506	TGCAAGTGGGCTGGAAGACATCCTGAAAGATGTCTAATTTTCAATCAAGGTACTTAAG	565	Db	1475	GTCAATGTCAGAAGATGGCCAGTACGATGATGTTTCTTGGAAACAGACATTTGGAATCTGTC	1534
Db	455	TGTAATTAAGTGGGAAGATGCCAAATACAGAAATGTGCAAAATTTTATCAGAGTACTTAG	514	QY	1574	CTTAAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTTAGAAGAGGTCTCTGCTGAA	1633
QY	566	GCATATAATCAGACTCACCTTGTAGCCTGTGGAACGGGGCTTTTTCATCCAAATTTGCACC	625	Db	1535	CTCAAGTTGTGACGATTTCAAGGAAAGTGG---AATATGGAAGAGGTAGTGTCTGGAG	1591
Db	515	CCCTATAACAAAACCTCACATATATGTGTGGAACCTGGAGCATTTTCATCCAAATATGGG	574	QY	1634	GAAATGACAGTTTTCGGGAAACCGACTGCTATTTTTCAGCAATGGAGCTTTTCCATGAAGCAG	1693
QY	626	TACATTTGAAATTTGGACATCATCTCTGAGGACATATTTTAAAGTGGAGAACTCACATTTT	685	Db	1592	GAGTTGCAGATATTCAGCACTCATCAATCATCTTTGAACATGGAATTTGTCTCTGAAGCAG	1651
Db	575	TATATTTGATCTTTGGAGTCTACAAAGGAGATATTTATTTCAAACTAGACACATATTTTG	634	QY	1694	CAACAACTATATATTTGTTTCAACGCTGGGGTTCGCCAGCTCCCTTTTACACCGGTGTGAT	1753
QY	686	GAAACGGGCTGGGAAGAGTCCATATGACCTTAAGCTGTCGACAGCATCCCTTTTAAATA	745	Db	1652	CAACAAATTTGATTTGTTTCCCGAGATGGATTTAGTTTCTCTCTGACAGATGCGAC	1711
Db	635	GAGTCTGGGAGACTGAAATGTCTTTCGATCTCTCAGCAGCCTTTTGTCTCAGTAATGACA	694	QY	1754	ATTTACGGGAAGGCTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTCTTTGGGAT	1813
QY	746	GATGGAGAAATPATCTCTGGAATCGAGCTGATTTTATGGGCGAGACTTTTCTATCTTC	805	Db	1712	ACTTATGGGAAGCTTGGCAGACTGTTGTTCTTGGCAGAGACCTTACTGTGCTCTGGAT	1771
Db	695	GATGAGTACTCTCTCTGGAACAGCTTCTGATTTCTTTGGCAAGATAGTCAATCACT	754	QY	1814	GGTCTGCAATGTTCTGCTATTTTCCACTGGAAGAGACGCAAGAGAGCAAGATATA	1873
QY	806	CGAACTCTTGGGCTACTCATGACCACTCATCATCAAGAACTGACATTTTTCAGAGCACTAC	814	Db	1772	GGAAATGATGCTCTCGATATGCTCTACTTCTTAAAGGAGAGCTAGACGCCAAGATGTA	1831
Db	755	CGATCCTTTTGGGCTTACTCATGACCACTCATCATCAAGAACTGACATTTTTCAGAGCACTAC	814	QY	1874	AGAAATGAGAGACCACTGACCTGCTTTCAGACTTACACCATGATTAATCACCATGCCAC	1933
QY	857	TGCTCAATGATCCAAAGTTTCAATAGTGGCCACCTCATCTCAGAGAGTACAAATCTGAA	916	Db	1832	AAATATGGGACCCCAATCACCAGTGTGGGACATCGAAGACAGCATTTAGTCAAG---AA	1888
Db	815	TGCTCAATGGAGCAAAATTTATGGAATTTCTTCTCATACAGACACCTTACAAATCCAGAT	874	QY	1934	AGCCCTGAAGAGAGAAATCATCTATGGTGTAGAGAAATAGTAGCACATTTTGGAAATGCAGT	1993
QY	917	GATGACAAAGTATPATCTTTTCTTCGTGAAATGCAATAGATGGAGAACACTCTGGAATA	976	Db	1889	ACTGCTGATGAAGAGTGAATTTTGGCATTGAATTTAACTCAACCTTTCTGGAATGTATA	1948
Db	875	GATGATAAAATATATTTCTTTCTGTTGATCATCTCAAGAGGCAAGTACCTCCGATAAA	934	QY	1994	CCGAAGTCGAGAGAGCCCTGCTATTTGGCAATTCGAGAGCGGAAATGAAGAGCGAAA	2053
QY	977	GCTACTCAGCTAGAAATAGTTCAGATATGCAAGATGACTTTGGAGGGCAGAGAGTCTG	1036	Db	1949	CTTAAATCCCAACAGCAACTATTTAAATGGTATATCCAGAGGTCAAGGAGTGAAGCATCGA	2008
Db	935	ACCATCTTTCTCGAGTTGGAAGAGTTTGAAGATGATGTAGGAGGCAACGCGACCTG	994	QY	2054	GAAGAGATCAGATGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGTGT	2113
QY	1037	GTGAATAATGAGACATTTCTCAAGCTCGTCTGATTTGCTCAGTGCAGGTCCAAT	1096	Db	2009	GAGGAGTTTGAAGCCCGATGAAGAAATCATCAAAACGGAATATGGGCTTACTGATTCGAAGT	2068
Db	995	ATAACAAGTGGAGACTTTTCTTAAGGCCAGACTGATTTGCTCAATTTCTGGAAGTGT	1054	QY	2114	CTACAACAAGAGGATTCAGGCAATTTACCTCTGCCATGCGGTGGAAACATGGGTTTATACAA	2173
QY	1097	GGCATTGACACTCATTTTGTGAACTGCAAGATGATTTCTTAATGAACTTTTAAAGATCCT	1156	Db	2059	TTGCAGAAGAGGATTTCTGGGATGTATTTACTGCAAGGCCAGGAGACACATTTTCAATCCAC	2128
Db	1055	GGGCGAGATACTTACTTTGATAGCTTCAAGATATTTATTTTACTCCCAACAGAGATGAA	1114	QY	2174	ACTCTTTTAAAGTAAACCTTGGAGTCAATGACACAGAGCATTTTGGAGAGAACTTTCTCAT	2233
QY	1157	AAAAATCCAGTTGTATATGAGTGTATGACTTCCAGTAAATTTTCAAGGATCAGCC	1216	Db	2129	ACCATAGTGAAGCTGACTTTTGAATGTCATTGAGAATGAACAGATGGAATAATACCCAGAGG	2188
Db	1115	AGAAATCCCTGTATGATGAGTCTTTTACTACAAACAGCTCCATCTTCAAGGCTCTGCT	1174	QY	2234	AAAGTATGATGATG 2246	
QY	1217	GTGTGTATGTATGAGATGATGTGAGAGGCTTCTTGGTCCATATGCCACAGG	1276	Db	2189	GCAGAGCATGAGG 2201	
Db	1175	GTGTGTGTATGAGATGCTGACATCAGAGCAGTTTTTAAATGTTCCATATGCTCATAG	1234				
QY	1277	GATGGACCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTTATCCACGGCCAGGA	1336				

RESULT 13
ACA05873
ID ACA05873 standard; cDNA; 3871 BP.
XX
ACA05873;
XX

DT 29-MAY-2003 (first entry)
XX Human secreted/transmembrane protein (PRO) cDNA #174.
DE
XX
KW Human; gene; ss; secreted and transmembrane protein; PRO; TNF-alpha;
KW tumour necrosis factor alpha; chondrocyte cell; tumour; gene therapy;
KW tissue typing.
XX Homo sapiens.
OS
XX
XX US2003036162-A1.
XX
XX 20-FEB-2003.
XX
XX 12-JUL-2002; 2002US-0194423.
XX
XX 16-SEP-1998; 98WO-US19330.
XX 07-OCT-1998; 98WO-US21141.
XX 01-DEC-1998; 98WO-US25108.
XX 08-MAR-1999; 99WO-US05028.
XX 14-MAY-1999; 99WO-US10733.
XX 02-JUN-1999; 99WO-US12252.
XX 01-SEP-1999; 99WO-US20111.
XX 15-SEP-1999; 99WO-US21090.
XX 01-DEC-1999; 99WO-US28301.
XX 02-DEC-1999; 99WO-US28551.
XX 30-DEC-1999; 99WO-US31274.
XX 05-JAN-2000; 2000WO-US00219.
XX 18-FEB-2000; 2000WO-US04341.
XX 18-FEB-2000; 2000WO-US04342.
XX 22-FEB-2000; 2000WO-US04414.
XX 24-FEB-2000; 2000WO-US05004.
XX 01-MAR-2000; 2000WO-US05601.
XX 02-MAR-2000; 2000WO-US05841.
XX 15-MAR-2000; 2000WO-US06884.
XX 30-MAR-2000; 2000WO-US08439.
XX 17-MAY-2000; 2000WO-US13705.
XX 22-MAY-2000; 2000WO-US14042.
XX 30-MAY-2000; 2000WO-US14941.
XX 02-JUN-2000; 2000WO-US15264.
XX 28-JUL-2000; 2000WO-US20710.
XX 24-AUG-2000; 2000WO-US23328.
XX 08-NOV-2000; 2000WO-US30952.
XX 01-DEC-2000; 2000WO-US32678.
XX 20-DEC-2000; 2000WO-US34956.
XX 28-FEB-2001; 2001WO-US06520.
XX 01-JUN-2001; 2001WO-US17800.
XX 20-JUN-2001; 2001WO-US19692.
XX 29-JUN-2001; 2001WO-US21066.
XX 09-JUL-2001; 2001WO-US21735.
XX 29-AUG-2001; 2001WO-US27099.
XX 26-JUN-1998; 98US-0105413.
XX 07-OCT-1998; 98US-0168978.
XX 06-NOV-1998; 98US-0187368.
XX 07-DEC-1998; 98US-0202054.
XX 03-MAR-1999; 99US-0254311.
XX 14-MAY-1999; 99US-0311832.
XX 14-MAY-1999; 99US-0380137.
XX 25-AUG-1999; 99US-0380138.
XX 25-AUG-1999; 99US-0380139.
XX 25-AUG-1999; 99US-0380142.
XX 18-OCT-1999; 99US-0403297.
XX 12-NOV-1999; 99US-0423844.
XX 28-AUG-2000; 2000US-0644848.
XX 18-SEP-2000; 2000US-0646410.
XX 18-SEP-2000; 2000US-0665350.
XX 08-NOV-2000; 2000US-0709238.
XX 20-DEC-2000; 2000US-0747259.
XX 22-MAR-2001; 2001US-0816744.
XX 10-MAY-2001; 2001US-0854208.
XX 10-MAY-2001; 2001US-0854280.
XX 25-MAY-2001; 2001US-0866028.
XX 05-JUN-2001; 2001US-0874503.
PR 18-JUL-2001; 2001US-0908827.
PR 30-JUL-2001; 2001US-0918585.
PR 06-AUG-2001; 2001US-0924419.
PR 13-AUG-2001; 2001US-0929404.
PR 16-AUG-2001; 2001US-0931836.
PR 28-AUG-2001; 2001US-0941992.
PR 04-SEP-2001; 2001US-0946374.
PR 15-JAN-2002; 2002US-0052586.
XX
XX (GETH) GENENTECH INC.
XX
XX Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
XX Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
XX WPI; 2003-332039/31.
XX P-PSDB; ABU67579.
XX
XX New secreted and transmembrane PRO polypeptides and nucleic acids,
XX useful in gene therapy, in chromosome and gene mapping, as chromosome
XX markers, in tissue typing, and in chromosome identification -
XX
XX Claim 2; Fig 347; 706pp; English.
XX
XX The invention discloses human nucleic acids encoding secreted and
XX transmembrane (PRO) polypeptides. Also disclosed is an antibody that
XX specifically binds to the PRO polypeptide, a method for stimulating the
XX release of tumour necrosis factor alpha (TNF-alpha) from human blood by
XX contacting the blood a PRO polypeptide, a method for stimulating the
XX proliferation or differentiation of chondrocyte cells by contacting the
XX cells with a PRO polypeptide, a method for detecting the presence of a
XX tumour in a mammal and an oligonucleotide probe derived from any of the
XX PRO nucleotide sequences. The nucleotide sequences are useful as probes,
XX in chromosome and gene mapping, in generating antisense RNA and DNA, in
XX preparing PRO polypeptides by recombinant techniques and in gene therapy
XX (e.g. for replacement of defective gene). The PRO polypeptides are useful
XX as molecular weight markers for protein electrophoresis purposes, for
XX chromosome identification, as chromosome markers, as therapeutic agents,
XX for stimulating the release of TNF-alpha from human blood, for
XX stimulating the proliferation or differentiation of chondrocytes and
XX detecting the presence of a tumour. The PRO polypeptides and nucleic
XX acids may also be used diagnostically for tissue typing. The sequences
XX presented in ACA05700-ACA06004 are the cDNAs encoding the PRO
XX polypeptides of the invention.
SQ Sequence 3871 BP; 1251 A; 704 C; 805 G; 1111 T; 0 other;
Query Match 26.5%; Score 718.6; DB 25; Length 3871;
Best Local Similarity 61.9%; Pred. No. 4.2e-206;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;
QY 269 AATGGGAAGAACAAATGTGCCAAGGCTGAAATTTATCTTACAAAGAAATGTTGGAATCCAAAC 328
DB 215 ACTTTGAAGCAAAATATTCAGACTCAAGTACTTACAAAGACTTGCTGCTTCAAT 274
QY 329 AATGTGATCACTTTCAATGGCTTGGCCAAAGCTCCAGTTATCATACCTTCCTTTGGAT 388
DB 275 AGCTGTATTCCTTTTGGGTTTCATCAGAGGACTGATTTTCAACTCTTCTCTAGAT 334
QY 389 GAGGAACGGAGTAGGCTGTATGTTGGAGCAAGAGTACATATTTTTCATTCGACTGGTT 448
DB 335 GAGGAAGAGCGCAGGCTGCTTTGGGAGCCAAAGACCAATCTTTTCTACTCAGTCTGGTT 394
QY 449 AA---TATCAAGGATTTTCAAAAGATTGTGTGGCCAGTATCTTACACAGAAAGAGATGAA 505
DB 395 GACTTAAACAAAATTTTAAAGAGATTATTTGGCTGCTGCAAGGAGCGGTGAATTA 454
QY 506 TGCAAGTGGGCTGAAAGACATCTCTGAAAGAAATGTGCTAATTTTCATCAAGGTACTTTAAG 565
DB 455 TGTAAATTTAGCTGGGAAGATGCCAATACAGAGATGTGCAAAATTTTCATCAGAGTACTTCAG 514
QY 566 GCATATAATCAGACTCACTTGTAGCCCTGTGGACGGGGGCTTTTCATCCAAATTTTGACC 625
DB 515 CCTATAACAAACCTCATATATGTGTGTGGAACTGGAGCATTTTTCATCAATATGTGGG 574

QY 626 TACATTGAAATGGACATCATCTGAGGCAATATATTTTAAAGCTGGAGAACTCACATTTT 685
DB 575 TATATTGATCTTGGAGTCTCAAGGAGGATATATATTCAACTAGACACATATTTG 634
QY 686 GAAACCGCGGTGGGAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAAATA 745
DB 635 GAGTCTGGCAGAGCTGAAATGTCTCTTCGATCTCTCAGCAGCCTTTTTCAGTAAATGACA 694
QY 746 GATGGAGATATATCTCTGGAATCTGAGCTGATTTATGGGGGAGACCTTTGCTATCTTC 805
DB 695 GATGAGTACCTCTACTCTGGAACAGCTTCTGATTTCTTTGGCAAGATATCTGCATTCAT 754
QY 806 CGAATCTTTGGGC-----ACCACCCCAATCAGGACAGAGCAGCATGATCCAGG 856
DB 755 CGATCCCTTTGGGCTACTCATGACCACTACTATCAGAACTGACATTTTCAGAGCACTAC 814
QY 857 TGCTCAATGATCCAAAGTTCAATAGTGGCCACCTCATCTCAGAGAGTGAACAATCTGAA 916
DB 815 TGCTCAATGAGCAAAATTTATTTGAACTTTCTTCTCATACCAAGACACCTTACAATCCAGAT 874
QY 917 GATGACAAAGTATATCTTTCTTCGTTGAAATGCAATAGATGGAGAACACTCTGGAATA 976
DB 875 GATGATAAATATATTTCTTCTTCTGTAATCATCTCAAGAGGCGAGTACCTCCGATAAA 934
QY 977 GCTACTCACGCTAGATAGTCAAGATAGTCAAGATGACTTTGGAGGGCACAGAACTCTG 1036
DB 935 ACCATCTTCTCGAGTGGAGAGTTTGTAGATGATGTAGGAGGACACGAGCCCTG 994
QY 1037 GTGAATPAATGGACAACTTCTCAAAGCTCGTCTGATTTGCTCAGTGGCAGGTCCAAT 1096
DB 995 ATAAACAAGTGGAGAGCTTTCTTAAAGGCGAGACTGATTTGCTCAATCTCTGGAAGTAT 1054
QY 1097 GGCATTTGACACTCATTTTCTGATGACTGAGGATGATTTCTTAATGAACTTTAAAGATCT 1156
DB 1055 GGGCGAGATACCTTCTGATGAGCTTCAAGATATTTATTTACTCCCCCAAGAGATGAA 1114
QY 1157 AAAATCCAGTTGTATATGAGTGTATGACCTTCCAGTAACTTTTCAAGGGATCAGCC 1216
DB 1115 AGAATCTGTAGTATATGAGTCTTTACTACACAGGCTCCATCTTCAAGGCTCTGCT 1174
QY 1217 GTGTGATGATAGCATAGTATGATGAGAAGGGTGTCTTTGGTCCATATGCCACAGG 1276
DB 1175 GTTTGTGTATAGCATGCTGACATCAGAGCAGTTTTTAAATGGTCCATATGCTCATAAG 1234
QY 1277 GATGGACCAACTATCAATGGTGTCTTATCAAGGAGAGTCCCTATCCAGCGGCAGGA 1336
DB 1235 GAAAGTGCAGACCATCGTTGGGTGAGTATGATGGGAGAAATTCCTTATCCACGGCTGGT 1294
QY 1337 ACTTGTCCAGCAAAACATTTG---GTGGTGTGACTCTCAAAAGGACCTTCTCTGATGAT 1393
DB 1295 ACATGTCGAGCAAAACCTATGACCCACTGATTAAGTCCACCGAGATTTTCCAGATGAT 1354
QY 1394 GTTATAACCTTTGCAAGAAGTATCCAGCCATGATCAATCCAGTGTTCCTATGAACAAT 1453
DB 1355 GTCATCAGTTTCATAAAGCGGCACTCTGTGATGATAGTAAAGTCCGTATACCCAGTTGCAGGA 1414
QY 1454 CGCCCAATAGTATCAAAACGGATGTAANTTATCAATTTACACAAATGTCGTAGACCGA 1513
DB 1415 GGACCAACGTTCAAGAGAAATCAATGTGGATTAACAGACTGACACAGATAGTGTGGATCAT 1474
QY 1514 GTGGATGCAGAGATGGACAGTATGATGTTATGTTATCGGAACAGATGTTGGGACCGTT 1573
DB 1475 GTCAATGCAAGATGGCCAGTACGATGTAATGTTTCTTGGAAACAGACATTTGGAAGTCTC 1534
QY 1574 CTTAAAGTATGTTCAATCTTAAAGGAGACTTGGTATGATTTAGAAAGAGTCTCTGCTGAA 1633
DB 1535 CTTAAAGTGTGAGCATTTTCAAAGGAAAGTGG---AATATGGAAGAGGTAGTGTCTGGAG 1591
QY 1634 GAAATGACAGTTTCTTGGGAAACGAGTCTGATTTTTCAGCAATGGAGCTTTCACACTAAGCAG 1693
DB 1592 GAGTTGAGATATTTCAAGCACTCATCAATCATCTTGTAAACATGGAATTTGTCTCTGAAGCAG 1651

RESULT 14

ABX97917

ID ABX97917 standard; cDNA; 3871 BP.

XX AC ABX97917;

XX XX 16-MAY-2003 (first entry)

XX Human PRO polynucleotide #174.

XX Human; PRO; gene; ss; cytostatic; chromosome mapping; gene mapping;
XX protein electrophoresis; tumour necrosis factor-alpha; TNF-alpha; blood;
XX chondrocyte differentiation; chondrocyte proliferation; tumour.

XX OS Homo sapiens.

XX FN US2003032102-A1.

XX PD 13-FEB-2003.

XX PF 17-JUN-2002; 2002US-0173697.

XX PR 16-SEP-1998; 98WO-US19330.

XX PR 07-OCT-1998; 98WO-US21141.

XX PR 01-DEC-1998; 98WO-US25108.

XX PR 08-MAR-1999; 99WO-US05028.

XX PR 14-MAY-1999; 99WO-US10733.

XX PR 02-JUN-1999; 99WO-US12252.

XX PR 01-SEP-1999; 99WO-US20111.

XX PR 15-SEP-1999; 99WO-US21090.

XX PR 01-DEC-1999; 99WO-US28301.

XX PR 02-DEC-1999; 99WO-US28551.

PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04341.
PR 22-FEB-2000; 2000WO-US04342.
PR 24-FEB-2000; 2000WO-US04414.
PR 01-MAR-2000; 2000WO-US05004.
PR 02-MAR-2000; 2000WO-US05601.
PR 15-MAR-2000; 2000WO-US05941.
PR 30-MAR-2000; 2000WO-US06984.
PR 17-MAY-2000; 2000WO-US08439.
PR 22-MAY-2000; 2000WO-US13705.
PR 30-MAY-2000; 2000WO-US14042.
PR 02-JUN-2000; 2000WO-US14941.
PR 28-JUL-2000; 2000WO-US15264.
PR 24-AUG-2000; 2000WO-US20710.
PR 08-NOV-2000; 2000WO-US23328.
PR 01-DEC-2000; 2000WO-US30952.
PR 20-DEC-2000; 2000WO-US32678.
PR 28-FEB-2001; 2000WO-US32956.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 29-AUG-2001; 2001WO-US27099.
PR 18-SEP-1997; 97US-059263P.
PR 18-SEP-1997; 97US-059266P.
PR 17-OCT-1997; 97US-062250P.
PR 21-OCT-1997; 97US-063486P.
PR 24-OCT-1997; 97US-063120P.
PR 24-OCT-1997; 97US-063121P.
PR 28-OCT-1997; 97US-063340P.
PR 28-OCT-1997; 97US-063541P.
PR 28-OCT-1997; 97US-063544P.
PR 28-OCT-1997; 97US-063564P.
PR 29-OCT-1997; 97US-063734P.
PR 31-OCT-1997; 97US-063870P.
PR 31-OCT-1997; 97US-064103P.
PR 13-NOV-1997; 97US-065311P.
PR 21-NOV-1997; 97US-066120P.
PR 24-NOV-1997; 97US-066466P.
PR 24-NOV-1997; 97US-066722P.
PR 11-DEC-1997; 97US-069335P.
PR 12-DEC-1997; 97US-069425P.
PR 17-DEC-1997; 97US-069870P.
PR 18-DEC-1997; 97US-068017P.
PR 10-MAR-1998; 98US-077450P.
PR 11-MAR-1998; 98US-077632P.
PR 11-MAR-1998; 98US-077849P.
PR 20-MAR-1998; 98US-078886P.
PR 20-MAR-1998; 98US-078939P.
PR 27-MAR-1998; 98US-079664P.
PR 27-MAR-1998; 98US-079786P.
PR 31-MAR-1998; 98US-080107P.
PR 31-MAR-1998; 98US-080194P.
PR 01-APR-1998; 98US-080327P.
PR 01-APR-1998; 98US-080333P.
PR 08-APR-1998; 98US-081049P.
PR 08-APR-1998; 98US-081070P.
PR 09-APR-1998; 98US-081195P.
PR 15-APR-1998; 98US-081838P.
PR 21-APR-1998; 98US-082568P.
PR 21-APR-1998; 98US-082569P.
PR 22-APR-1998; 98US-082704P.
PR 22-APR-1998; 98US-082797P.
PR 28-APR-1998; 98US-083222P.
PR 29-APR-1998; 98US-083495P.
PR 29-APR-1998; 98US-083496P.
PR 29-APR-1998; 98US-083499P.
PR 29-APR-1998; 98US-083559P.
PR 05-MAY-1998; 98US-084366P.
PR 06-MAY-1998; 98US-084414P.
PR 07-MAY-1998; 98US-084639P.

PR 07-MAY-1998; 98US-084640P.
PR 07-MAY-1998; 98US-084643P.
PR 15-MAY-1998; 98US-085579P.
PR 15-MAY-1998; 98US-085580P.
PR 15-MAY-1998; 98US-085582P.
PR 15-MAY-1998; 98US-085700P.
PR 18-MAY-1998; 98US-086023P.
PR 22-MAY-1998; 98US-086392P.
PR 22-MAY-1998; 98US-086486P.
PR 28-MAY-1998; 98US-087098P.
PR 28-MAY-1998; 98US-087208P.
PR 02-JUN-1998; 98US-087609P.
PR 02-JUN-1998; 98US-087759P.
PR 03-JUN-1998; 98US-087827P.
PR 04-JUN-1998; 98US-088025P.
PR 04-JUN-1998; 98US-088028P.
PR 04-JUN-1998; 98US-088029P.
PR 04-JUN-1998; 98US-088033P.
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.

PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-096959P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 01-SEP-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098116P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.
PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.
PR 10-SEP-1998; 98US-099754P.
PR 10-SEP-1998; 98US-099763P.
PR 10-SEP-1998; 98US-099812P.
Query Match 26.5%; Score 718.6; DB 25; Length 3871;
Best Local Similarity 61.9%; Pred. No. 4.2e-206; Mismatches 739; Indels 21; Gaps 5;
Matches 1233; Conservative 0;

QY 269 AATGGAAGAACAAATGTGCCAAGGCTGAAATATCTCAACAAAGAAATGTGGAATCCAAC 328
DB 215 ACTTTGAAGCAAAATATCCAGACTCAAGCTAACTCAAGAGACTTGTGCTTCAAT 274
QY 329 AATGTGATCATTCAATGGCTTGGCCAAAGCTCAGTTCATACCTTCTTTTGGAT 388
DB 275 AGCTGTATTCCTTTTGGGTTTCATCAGAAGGACTGGATTTTCAAACTCTTCTTAGAT 334
QY 389 GAGGAACGAGTAGGCTGTATGTGGAGCAAGAGATCAGATTTTTCATTCGACCTGGTT 448
DB 335 GAGGAAGAGGAGGCTGCTCTTTGGAGCCAAAGACCAATCTTCTACTCAGTCTGGTT 394
QY 449 AA---TATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACACCAAGAGAGATGAA 505
DB 395 GACTTAAACAAAAATTTTAAAGAGATTTATTTGGCTGTGCTGCAAGAGAGGCTGGAATTA 454
QY 506 TCGAAGTGGCTGGAAAGACATCTCGAAAGAAATGTGCTTAATTTCAATCAAGTACTTAAG 565
DB 455 TGTAAATTAGCTGGGAAGATGCCAATACAGAAATGTGCAAAATTTTCATCAGAGTACTTCAG 514
QY 566 GCATATAATCAGACTCAGCTTGTAGGCTGTGGAAGGGGGCTTTTCATCCAAATTTGCACC 625
DB 515 CCTATTAACAAACTCACATATATGTGTGGAAGTGTGGAAGTGTGGAAGTGTGGAAGTGTGGA 574
QY 626 TACATTGAAATTTGACATCATCTGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTT 685
DB 575 TATATTGATCTTGGAGTCTACAAGGAGGATATATATTCAAACTAGACACACATATTG 634
QY 686 GAAACGGCGGTGGGAAGAGTCCATATGACCTTAAGTCTGCTGACAGCATCCCTTTTAATA 745
DB 635 GAGTCTGGCAGACTGAAATGTCTTTCGATCTCTCAGCAGCCTTTTGTCTTCAATATGACA 694
QY 746 GATGAGAAATTAATCTCTGGAAGTCAAGTGTGGAAGGGGGCTTTTATGGGGGAGACTTTGCTATCTTC 805
DB 695 GATGAGTACTCTACTCTGGAACAGCTTCTGATTTCTTGGCAAGATCTGATTTCACT 754
QY 806 CGAACTCTTTGGGC-----ACCACCAACCAATCAGGACAGAGAGCATGATTTCCAGG 856
DB 755 CGATCCTTTGGGCTTACTCATGACCACTACATACAGAACTGACATTTTCAGAGCACTAC 814
QY 857 TGGCTCAATGATCCAAAGTTTCAATAGTGGCCACCTCATCTCAGAGAGTACAAATCCTGAA 916
DB 815 TGGCTCAATGGAGCAAAATTTTATGGAACTTTTCTTATACCAAGCAGCCTTACAAATCCAGAT 874

QY 917 GATGACAAAGTATACTTTTCTTCGTCGAAATGCAATAGATGGAGAACTCTCTGAAAA 976
DB 875 GATGATAAAATATATTTTCTTTCTGTCGAAATCATCTCAAGAGGAGTACCTCCGATATA 934
QY 977 GCTACTCAGCTAGGATAGGTCAGATATGCAAGAAATGACTTTGGAGGGCACAAGATCTG 1036
DB 935 ACCATCTCTCTCGAGTTGGAAGAGTTGTGAAGATGATGTAGGAGGACAACGACGCTG 994
QY 1037 GTGAATAATGACAAACATTTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCCTCAAT 1096
DB 995 ATAAACAAGTGGAGCACTTTTCTTAAGGCCAGAGCTGATTTGCTCAATTTCTTGAAGTAT 1054
QY 1097 GGCATTGACACTCAATTTTGTGAACCTGCAAGATGATTTCTTAATGAACTTTTAAAGATCCT 1156
DB 1055 GGGGAGATATCTTACTTTGATGAGCTTCAAGATATTTATTACTCCCAAGAGATGAA 1114
QY 1157 AAAAATCAGTTGTATATGGAGTGTTTACGATCTTCAGTAAACATTTTCAAGGATCAGGC 1216
DB 1115 AGAAATCCTGTAGTATATGGAGTCTTTACTACAACAGCTCCATCTTCAAGGCTCTGCT 1174
QY 1217 GTGTGTATATAGCATGATGTGAGAGGGTGTTCCTTGGTCCATATGCCACACAGG 1276
DB 1175 GTTTGTGTATAGCATGGCTGACATCAGAGCAGTTTTAAATGGTCCATATGCTCATAG 1234
QY 1277 GATGACCAACTATCAATGGTGTCTTATCAAGGAAGAGTCCCCTATCCAGCGCCAGGA 1336
DB 1235 GAAAGTCAGACCACTGTGTGGTGCAGTATGATGGAGAAATCTTATCCAGGCTGGT 1294
QY 1337 ACTTGTCCCAACAAAACATTTG---GTGGTTTGTACTCTCAAAAGGACCTTCTCTGATGAT 1393
DB 1295 ACATGTCCAAGCAAAACCTATGACCCACTGATTAAGTCCACCCGAGATTTTCCAGATGAT 1354
QY 1394 GTTATAACCTTTGCAAGAGTCATCCAGCATGATGATCAATCCAGTGTTCCTTATGAACAA 1453
DB 1355 GTCATCAGTTTCATAAAGCGGCACTCTGTGATATATAAGTCGGTATATCCAGTTTCAGGA 1414
QY 1454 CGCCCAATAGTGTATCAAAACGGATGTAATTTATCAATTTACACAATTTGTGTAGACCGA 1513
DB 1415 GGACCAAGTTCAAGAGATCAATGTGGNTTACAGACTGACACAGATAGTGTGATCAT 1474
QY 1514 GTGATGACAGAGATGACAGTATGATGATTTATGTTTATCGGAACAGATGTTGGGACCGTT 1573
DB 1475 GTCATGACAGAGATGCGCAGTACGATGATTAATTTCTTGGAAACAGACATTTGGAACCTGTC 1534
QY 1574 CTTAAAGTGTTCATTTCTTAAGGAGACTTGGTATGATTTTAGAAGAGTGTCTGCTGGA 1633
DB 1535 CTCAAAGTGTGAGCATTTTCAAGGAAAGTGG---AATATGGAAGAGGTAGTGTCTGGAG 1591
QY 1634 GAAATGACAGTTTTTTCGGGAACCGACTGCTATTTTCAAGCAATGGAGCTTTCCACTAAGCAG 1693
DB 1592 GAGTTGCAGATATTTCAAGCACTCATCAATCATCTTGAACATGGAATTTGCTCTGAGCAG 1651
QY 1694 CAACAACATATATTTGGTTCAACGGCTGGGGTGGCCAGTCCCTTTTACACCGGTGTGAT 1753
DB 1652 CAACAATTTGATACATTTGGTTCCCGAGATGGAATAGTTAGTTAGTCTCTCTTGCACAGATGCGAC 1711
QY 1754 ATTTACGGGAAGCGTGTGCTGAGTGTGCTCTCCCGAGAGCCCTTACTGCTGTTGGAT 1813
DB 1712 ACTTATGGGAAGCTTGGCAGACTGTTGTCTTCCAGAGACCCCTACTGTGCTCTGGGAT 1771
QY 1814 GGTTCGTGATCTCTCGCTATTTTCCCACTGCAAGAGACGCAAGACGCAAGATATA 1873
DB 1772 GGAATGATGCTCTCGATATGCTCTACTTCTTAAGGAGAGCTAGACGCCAGATGTA 1831
QY 1874 AGAAATGGAGACCACTGACTGCTGTTTACAGCTTTACACCATGATATAATCACTGGCCAC 1933
DB 1832 AAATATGCGACCCCAATCACCAGTGTGCGACATCGAAGACAGCATTTAGTCAATG---AA 1888
QY 1934 AGCCCTGAGAGAGAAATCATCTATGTTGTGATAGAAATAGTACACATTTTTCGAATGCAAT 1993
DB 1889 ACTGCTGATGAAAGGAGTATTTTGGCAATGAAATTTTAACTCAACCTTTCTGGAATGTATA 1948
QY 1994 CCGAAGTCGAGAGAGCGCTGCTCTATTGGAATTTCCAGAGCGGCAATTAAGAGCGCAAAA 2053

```

Db      1949  CTTAAATCCCAACAGCAACTATTAAATGGTATATCCAGAGTTCAGGGATGAGCATCGA 2008
Qy      2054  GAAGAGATCAGAGTGGATGATCATATCATCAGACAGATCAAGGCCCTTCTGCTAGCTAGT 2113
Db      2009  GAGGAGTTGAGCCCGATGAAGATCATCAAAACGGAATATGGCTACTGATTGGAAGT 2068
Qy      2114  CTACACAGAGAGGATTCAGGCAATTACCTCTGCCATGCGGTGGAACATGGGTTTCATACAA 2173
Db      2069  TTGCAGAGAGAGGATTTCTGGGATGTATTACTGCAAAAGCCCGAGGACACACTTTTCATCCAC 2128
Qy      2174  ACTCTCTTTAAGSTACCTGGAAGTCATTGACACAGAGCAATTTGGAAGAACTTTCTTCAT 2233
Db      2129  ACCATAGTGAAGCTGACTTTGAATGTTCATTGAGAATGAACAGATGGAATAATACCCAGAGG 2188
Qy      2234  AAAGATGATGATG 2246
Db      2189  GCAGAGCATGAGG 2201

RESULT 15
ID  ABX78701 standard; cDNA; 3871 BP.
AC  ABX78701;
XX
XX
DT  15-APR-2003 (first entry)
DE  Human PRO polynucleotide #174.
KW  Human; PRO; gene; ss; cytostatic; tumour; cancer; breast; lung; stomach;
KW  liver; dog; cat; cow; horse; sheep; pig; goat; rabbit; ADEPT;
KW  antibody-dependent enzyme mediated prodrug therapy.
OS  Homo sapiens.
PN  US2003027272-A1.
XX
XX
PD  06-FEB-2003.
XX
XX
PF  21-JUN-2002; 2002US-0176492.
XX
XX  16-SEP-1998; 98WO-US19330.
XX  07-OCT-1998; 98WO-US21141.
XX  01-DEC-1998; 98WO-US25108.
XX  08-MAR-1999; 99WO-US05028.
XX  10-MAR-1999; 99WO-US05190.
XX  14-MAY-1999; 99WO-US10733.
XX  02-JUN-1999; 99WO-US12252.
XX  01-SEP-1999; 99WO-US20111.
XX  15-SEP-1999; 99WO-US21090.
XX  30-NOV-1999; 99WO-US28313.
XX  01-DEC-1999; 99WO-US28301.
XX  02-DEC-1999; 99WO-US28551.
XX  30-DEC-1999; 99WO-US31274.
XX  05-JAN-2000; 2000WO-US00219.
XX  18-FEB-2000; 2000WO-US04341.
XX  18-FEB-2000; 2000WO-US04342.
XX  22-FEB-2000; 2000WO-US04414.
XX  24-FEB-2000; 2000WO-US05004.
XX  01-MAR-2000; 2000WO-US05601.
XX  02-MAR-2000; 2000WO-US05841.
XX  10-MAR-2000; 2000WO-US06319.
XX  15-MAR-2000; 2000WO-US06884.
XX  30-MAR-2000; 2000WO-US08439.
XX  17-MAY-2000; 2000WO-US13705.
XX  22-MAY-2000; 2000WO-US14042.
XX  30-MAY-2000; 2000WO-US14941.
XX  02-JUN-2000; 2000WO-US15264.
XX  28-JUL-2000; 2000WO-US20710.
XX  24-AUG-2000; 2000WO-US23328.
XX  08-NOV-2000; 2000WO-US30952.
XX  10-NOV-2000; 2000WO-US30873.

PR  01-DEC-2000; 2000WO-US32678.
PR  20-DEC-2000; 2000WO-US34956.
PR  28-FEB-2001; 2001WO-US06520.
PR  01-JUN-2001; 2001WO-US17800.
PR  20-JUN-2001; 2001WO-US19692.
PR  29-JUN-2001; 2001WO-US21066.
PR  09-JUL-2001; 2001WO-US21735.
PR  18-SEP-1997; 97US-059263P.
PR  18-SEP-1997; 97US-059266P.
PR  17-OCT-1997; 97US-062250P.
PR  21-OCT-1997; 97US-063486P.
PR  24-OCT-1997; 97US-063120P.
PR  24-OCT-1997; 97US-063121P.
PR  28-OCT-1997; 97US-063540P.
PR  28-OCT-1997; 97US-063541P.
PR  28-OCT-1997; 97US-063544P.
PR  28-OCT-1997; 97US-063564P.
PR  29-OCT-1997; 97US-063734P.
PR  31-OCT-1997; 97US-063870P.
PR  31-OCT-1997; 97US-064103P.
PR  13-NOV-1997; 97US-065311P.
PR  21-NOV-1997; 97US-066120P.
PR  24-NOV-1997; 97US-066466P.
PR  24-NOV-1997; 97US-066772P.
PR  11-DEC-1997; 97US-069335P.
PR  12-DEC-1997; 97US-069425P.
PR  17-DEC-1997; 97US-069870P.
PR  18-DEC-1997; 97US-068017P.
PR  10-MAR-1998; 98US-077450P.
PR  11-MAR-1998; 98US-077632P.
PR  11-MAR-1998; 98US-077649P.
PR  20-MAR-1998; 98US-078886P.
PR  20-MAR-1998; 98US-078939P.
PR  27-MAR-1998; 98US-079664P.
PR  27-MAR-1998; 98US-079786P.
PR  31-MAR-1998; 98US-080107P.
PR  31-MAR-1998; 98US-080194P.
PR  01-APR-1998; 98US-080327P.
PR  01-APR-1998; 98US-080333P.
PR  08-APR-1998; 98US-081049P.
PR  08-APR-1998; 98US-081070P.
PR  09-APR-1998; 98US-081195P.
PR  15-APR-1998; 98US-081838P.
PR  21-APR-1998; 98US-082568P.
PR  21-APR-1998; 98US-082569P.
PR  22-APR-1998; 98US-082704P.
PR  22-APR-1998; 98US-082797P.
PR  28-APR-1998; 98US-083322P.
PR  29-APR-1998; 98US-083495P.
PR  29-APR-1998; 98US-083496P.
PR  29-APR-1998; 98US-083499P.
PR  29-APR-1998; 98US-083559P.
PR  05-MAY-1998; 98US-084366P.
PR  06-MAY-1998; 98US-084414P.
PR  07-MAY-1998; 98US-084639P.
PR  07-MAY-1998; 98US-084640P.
PR  07-MAY-1998; 98US-084643P.
PR  15-MAY-1998; 98US-085579P.
PR  15-MAY-1998; 98US-085580P.
PR  15-MAY-1998; 98US-085582P.
PR  15-MAY-1998; 98US-085700P.
PR  22-MAY-1998; 98US-086023P.
PR  22-MAY-1998; 98US-086392P.
PR  22-MAY-1998; 98US-086486P.
PR  28-MAY-1998; 98US-087098P.
PR  28-MAY-1998; 98US-087208P.
PR  02-JUN-1998; 98US-087609P.
PR  02-JUN-1998; 98US-087759P.
PR  03-JUN-1998; 98US-087827P.
PR  04-JUN-1998; 98US-088025P.
PR  04-JUN-1998; 98US-088028P.
PR  04-JUN-1998; 98US-088029P.
PR  04-JUN-1998; 98US-088033P.
```

```
PR 04-JUN-1998; 98US-088326P.
PR 05-JUN-1998; 98US-088167P.
PR 05-JUN-1998; 98US-088202P.
PR 05-JUN-1998; 98US-088212P.
PR 05-JUN-1998; 98US-088217P.
PR 09-JUN-1998; 98US-088655P.
PR 10-JUN-1998; 98US-088722P.
PR 10-JUN-1998; 98US-088738P.
PR 10-JUN-1998; 98US-088740P.
PR 10-JUN-1998; 98US-088811P.
PR 10-JUN-1998; 98US-088824P.
PR 10-JUN-1998; 98US-088825P.
PR 10-JUN-1998; 98US-088826P.
PR 11-JUN-1998; 98US-088861P.
PR 11-JUN-1998; 98US-088863P.
PR 11-JUN-1998; 98US-088876P.
PR 12-JUN-1998; 98US-089090P.
PR 12-JUN-1998; 98US-089105P.
PR 16-JUN-1998; 98US-089512P.
PR 16-JUN-1998; 98US-089514P.
PR 17-JUN-1998; 98US-089538P.
PR 17-JUN-1998; 98US-089598P.
PR 17-JUN-1998; 98US-089653P.
PR 18-JUN-1998; 98US-089908P.
PR 19-JUN-1998; 98US-089952P.
PR 22-JUN-1998; 98US-090246P.
PR 22-JUN-1998; 98US-090252P.
PR 22-JUN-1998; 98US-090254P.
PR 24-JUN-1998; 98US-090429P.
PR 24-JUN-1998; 98US-090435P.
PR 24-JUN-1998; 98US-090444P.
PR 24-JUN-1998; 98US-090461P.
PR 24-JUN-1998; 98US-090535P.
PR 24-JUN-1998; 98US-090540P.
PR 25-JUN-1998; 98US-090676P.
PR 25-JUN-1998; 98US-090678P.
PR 25-JUN-1998; 98US-090688P.
PR 25-JUN-1998; 98US-090690P.
PR 25-JUN-1998; 98US-090694P.
PR 25-JUN-1998; 98US-090695P.
PR 25-JUN-1998; 98US-090696P.
PR 26-JUN-1998; 98US-090862P.
PR 26-JUN-1998; 98US-090863P.
PR 26-JUN-1998; 98US-091010P.
PR 01-JUL-1998; 98US-091359P.
PR 01-JUL-1998; 98US-091544P.
PR 02-JUL-1998; 98US-091478P.
PR 02-JUL-1998; 98US-091486P.
PR 02-JUL-1998; 98US-091626P.
PR 02-JUL-1998; 98US-091628P.
PR 02-JUL-1998; 98US-091632P.
PR 24-JUL-1998; 98US-094006P.
PR 04-AUG-1998; 98US-095282P.
PR 10-AUG-1998; 98US-095998P.
PR 10-AUG-1998; 98US-096012P.
PR 17-AUG-1998; 98US-096757P.
PR 17-AUG-1998; 98US-096766P.
PR 17-AUG-1998; 98US-096867P.
PR 17-AUG-1998; 98US-096891P.
PR 17-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096897P.
PR 18-AUG-1998; 98US-096949P.
PR 18-AUG-1998; 98US-097022P.
PR 26-AUG-1998; 98US-097952P.
PR 26-AUG-1998; 98US-097954P.
PR 26-AUG-1998; 98US-097955P.
PR 26-AUG-1998; 98US-097971P.
PR 26-AUG-1998; 98US-097974P.
PR 26-AUG-1998; 98US-098014P.
PR 01-SEP-1998; 98US-098716P.
PR 01-SEP-1998; 98US-098723P.
PR 02-SEP-1998; 98US-098803P.
PR 02-SEP-1998; 98US-098821P.

PR 02-SEP-1998; 98US-098843P.
PR 09-SEP-1998; 98US-099602P.
PR 10-SEP-1998; 98US-099741P.

Query Match 26.5%; Score 718.6; DB 25; Length 3871;
Best Local Similarity 61.9%; Pred. No. 4.2e-206;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;

QY 269 AATGGGAAGAACAAATGTCGCAAGGCTGAAATTTATCTCAAAAGAAATGTTGGAATCCAAC 328
DB 215 ACTTTGAAGCAAAATATTTCCAAGACTCAAGCTAACTCAAAAGACTTGTCTTCAAT 274
QY 329 AATGTGATCACTTTCAATGGCTTGGCCCAACAGCTCCAGTTATCATACATCTCTTTTGGAT 388
DB 275 AGCTGATTCCTTTTGGGTTTCATCAGAGACTGGATTTTCAAACTCTCTCTTAGAT 334
QY 389 GAGGAACGGAGTAGGCTGTATGTTGGAGCAAGGATCATATTTTCAATTCAGTCTGGTT 448
DB 335 GAGGAAGAGGCGAGCTGCTCTTGGGAGCCAAAGACCACATCTTCTACTCAGTCTGGTT 394
QY 449 AA---TATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACACCAGAGAGATGAA 505
DB 395 GACTTAAACAAAATTTTAAGAGATTTATTGGCCTGCTGCNAAGGAACGGGTGGAATTA 454
QY 506 TCGAAGTGGGCTGGAAAAGACATCTGAAAGAAATGCTAAATTTTCAATCAAGGTACTTAAG 565
DB 455 TGTAAATTTAGCTGGGAAAGATGCCAATACAGAAATGTCAAAATTTTCAATCAGAGTACTT 514
QY 566 GCATATAATCAGACTCACTTGTACGCTGTGGAAGCGGGGCTTTTCAATCCAAATTTGCACC 625
DB 515 CCTATAACAACCTCACATATATGTTGTGGAACCTGGAGCATTTTCAATCCAAATATGTGG 574
QY 626 TACATTGAAATTTGGACATCATCTGAGGACAAATATTTTAAAGCTGGAGAACTCACATTT 685
DB 575 TATATTGATCTGGAGTCTCAAGAGGATATTTATTTCAAACTAGACACACATATTTG 634
QY 686 GAAAACGGCCGTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAAATA 745
DB 635 GAGTCTGCGACACTGAAATGTCCTTTCGATCCTCAGACGCTTTTGTCTTCAATATGACA 694
QY 746 GATGAGAAATTTACTCTGGAACTGACGCTGATTTTATGGGGCGGAGACTTTGCTATCTTC 805
DB 695 GATGAGTACCTCTACTCTGGAACAGCTTCTGATTTCTTGGCAAGATATGTCATTCAT 754
QY 806 CGAATCTTTGGSC-----ACCACCACCAATCAGACAGAGCAGCATGATTTCCAGG 856
DB 755 CGATCCCTTTGGGCTTACTCATGACCCACCTACATCAGAACTGACATTTCAGAGCACTAC 814
QY 857 TGGCTCAATGATCCAAAGTTCAATTAGTCCCACTCATCTCAGAGAGTGACAACTCTGAA 916
DB 815 TGGCTCAATGAGCAAAATTTATTGGAATTTTCTTATACCAAGACACCTTCAATCCAGAT 874
QY 917 GATGCAAAAGTATATCTTTTCTTCCGTGAAATGCAATAGATGGAGAACACTCTCGAANA 976
DB 875 GATGATAAATATATTTCTTCTTCTGTAATCATCTCAAGGAAGGCGAGTACCTCCGATRAA 934
QY 977 GCTACTCAGCTAGAAATAGGTGAGATATGCAAGAAATGCAATAGATGGAGAACACTCTG 1036
DB 935 ACCATCTTCTCGAGTTGGAAGATTTGTAAGAAATGATAGGAGAACACGACGACCTG 994
QY 1037 GTGAATAATGGACAACATTCCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCCTCAAT 1096
DB 995 ATAAACAAGTGGACGACTTTTCTTAAAGCCAGACTGATTTGCTCAATTTCTTGAAGTGT 1054
QY 1097 GGCATTGACATCAATTTGATGAACTGCAGATGATTTCTTAATGAATTTTAAAGATCT 1156
DB 1055 GGGGCAGATACTTACTTTGATGAGCTTCAAGATATTTATTACTCTCCCAAGAGATGAA 1114
QY 1157 AAAAATCCAGTTGATATGGAGTGTTTACGACTTCCAGTAACATTTTCAAGGGATCAGCC 1216
DB 1115 AGAAATCCTGATGATATGGAGTCTTTACTAACCAGCTCATCTTCAAGGGCTCTGCT 1174
QY 1217 GTGTGATGATAGCATGAGTGATGTGAGAAGGGTGTCTTGTGTCCATATGCCACAGG 1276
```


Job time : 687 secs

```
Db 1175 GTTGTGTATAGCATGCTGACATCAGAGAGTTTTAAATGGTCCATATGCTCATAG 1234
Qy 1277 GATGACCCCAACTATCAATGGGTGCTTATCAAGAAAGAGTCCCTTATCCAGCGCAGGA 1336
Db 1235 GAAAGTGCAGACCATCGTTGGGTGCAGTATGATGGAGAAATTCCTTATCCAGCGCTGGT 1294
Qy 1337 ACTGTGCCAGCAAAACATTG---GTGGTTTGTACTTACAAAGAGCCCTCCTCATGAT 1393
Db 1295 ACATGTCCAAGCAAAACCTTATGACCCACTGATTAAGTCCACCCGAGATTTCCAGATGAT 1354
Qy 1394 GTTATAACCTTTGCAAGAGTCAATCCAGCATGTACATCCAGTGTTCCTATGAACAAT 1453
Db 1355 GTCATCAGTTTCATAAAGGGCACTCTGTGATGTATAAGTCCGTATACCCAGTTGCAGGA 1414
Qy 1454 CGCCCAATAGTATCAAAACGGATGTAAATTTCAATTTACACAAATTTGCTGTAGACCGA 1513
Db 1415 GGACCAACGTTCAAGAGATCAATGTGGATTACAGACTGCACACATAGTGGTGCATCAT 1474
Qy 1514 GTGGATGCAAGAGATGGACAGTATGATGTATGTATTCGGAACACAGATGTTGGACCGTT 1573
Db 1475 GTCAATTGCAGAGATGGCCAGTACGATGTAATGTTTCTTGGAAACAGACATTCGAACTGTC 1534
Qy 1574 CTTAAGTAGTTTCAATTCCTAAGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAA 1633
Db 1535 CTCAAAGTTGTGAGCATTTCAAGGAAAGTGG---AATATGGAAGAGGTAGTGTCTGGAG 1591
Qy 1634 GAAATGACAGTTTTTCGGAAACCGACTGCTATTTCCAGCAATGGAGCTTTCACATAAGCAG 1693
Db 1592 GAGTTGCAGATATTCAGACACTCATCATCTTGAACATGGAATTTGCTCTGAAGCAG 1651
Qy 1694 CAACAACTATATATTGGTTTCAACGGCTGGGGTTGCCAGCTCCCTTTTACACCGGTGTAT 1753
Db 1652 CAACAAATTGTACATTTGGTTCCCGAGATGGATTAGTTTCAGTCTCTCTTGCACAGATGCGAC 1711
Qy 1754 ATTTACGGGAAGCGTGTCTGAGTGTGCTCGCCCGAGACCCCTTACTGTGCTTGGAT 1813
Db 1712 ACTTATGGGAAAGCTTGGCGACATGTTGTCTTGGCCAGAGACCCCTTACTGTGCTGGAT 1771
Qy 1814 GGTCTGCTATGTTCTCGCTATTTTCCACTGCAAGAGAGCGCACAAAGACGACAGATATA 1873
Db 1772 GGAATGATGCTCTCGATATGCTCTTACTTCTAAAGAGAGCTAGACGCCAAGATGTA 1831
Qy 1874 AGAATGGAGACCCACTGACTCATCTGTTGAGACTTACACCATGATAATCACCATGGCCAC 1933
Db 1832 AAATATGGCGACCCAATCACCAGTGTGGGACATCGAAGACAGCATTAGTCATG---AA 1888
Qy 1934 AGCCCTGAAGAGAGATCATCTATGTTGTAGAGATAGTAGACACATTTTGGAAATGCAGT 1993
Db 1889 ACTGCTGATGAAAGGTGATTTTGGCATTGAATTTAACTCAACCTTTCTGGAATGTATA 1948
Qy 1994 CCGAAGTGCAGAGAGCGTGTCTATTGGCAATTCAGAGGGGAAATGAAGAGCGMAAA 2053
Db 1949 CCTAATCCCAACAGAACTAATTAATGGTATATCCAGAGGTGAGGGATGAGCATCGA 2008
Qy 2054 GAAGAGATCAGATGGATGATCATATCATCAGACAGATCAAGGCTTCTGCTACGTAGT 2113
Db 2009 GAGGAGTTGAAGCCCGATGAAAGATCATCAAAACGGAATATGGGTACTGATTCGAAGT 2068
Qy 2114 CTACAACAGAGAGATTACGGCAATTTACCTCTGCCATGGGTGGAAATGGGTTTCAACA 2173
Db 2069 TTGCAAGAAGAGGATTTCTGGGATGTATTACTGCAAAAGCCAGGAGCACACTTTTCATCCAC 2128
Qy 2174 ACTCTTCTTAAGGTAAACCTGGAAGTCAATTGGACACAGAGCATTTTGGAGAACTTCTTCAT 2233
Db 2129 ACCATAGTGAAGTCACTTTGAATGTCAATTGAGATGAACATGGAATAATACCCAGAGG 2188
Qy 2234 AAAGATGATGATG 2246
Db 2189 GCAGAGCATGAGG 2201
```

Search completed: July 31, 2003, 13:13:27

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 12:57:31 ; Search time 7224 Seconds
(without alignments)
12287.942 Million cell updates/sec

Title: US-09-774-490-1
Perfect score: 2709
Sequence: 1 aatctttatttcgatg.....aggcttttttctctaataacc 2709

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 32997241 seqs, 16383922548 residues

Total number of hits satisfying chosen parameters: 65994482

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents NA Main:

1: /cgn2_6/ptodata/1/pna/PCTUS COMB.seq.*
2: /cgn2_6/ptodata/1/pna/PCTUS COMB.seq.oid.*
3: /cgn2_6/ptodata/1/pna/US06 COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US07 COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US08 COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US081 COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US082 COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US083 COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US084 COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US085 COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US086 COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US087 COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US088 COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US089 COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US090 COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US091 COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US092 COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US092B COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US093A COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US093B COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US094 COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095A COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US095B COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US095C COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US095D COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096A COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096B COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US096C COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US096D COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US096E COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US097A COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US097B COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US097C COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US098A COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US098B COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US098C COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US098D COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US099A COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US099B COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US099C COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US099D COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US099E COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US099F COMB.seq.*

44: /cgn2_6/ptodata/1/pna/US100A COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US100B COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US101A COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US101B COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US102A COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US102B COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US103A COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US103B COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US104A COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US104B COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US6000 COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US6001 COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US6002 COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US6003 COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US6004 COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US6005 COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US6006 COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US6007 COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US6008 COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US6009 COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US6010 COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US6011 COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US6012 COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US6013 COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US6014 COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US6015 COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US6016 COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US6017 COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US6018 COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US6019 COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US6020 COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US6021 COMB.seq.*
76: /cgn2_6/ptodata/1/pna/US6022 COMB.seq.*
77: /cgn2_6/ptodata/1/pna/US6023A COMB.seq.*
78: /cgn2_6/ptodata/1/pna/US6023B COMB.seq.*
79: /cgn2_6/ptodata/1/pna/US6024 COMB.seq.*
80: /cgn2_6/ptodata/1/pna/US6025 COMB.seq.*
81: /cgn2_6/ptodata/1/pna/US6026 COMB.seq.*
82: /cgn2_6/ptodata/1/pna/US6027 COMB.seq.*
83: /cgn2_6/ptodata/1/pna/US6028 COMB.seq.*
84: /cgn2_6/ptodata/1/pna/US6029 COMB.seq.*
85: /cgn2_6/ptodata/1/pna/US6030 COMB.seq.*
86: /cgn2_6/ptodata/1/pna/US6031 COMB.seq.*
87: /cgn2_6/ptodata/1/pna/US6032 COMB.seq.*
88: /cgn2_6/ptodata/1/pna/US6033 COMB.seq.*
89: /cgn2_6/ptodata/1/pna/US6034 COMB.seq.*
90: /cgn2_6/ptodata/1/pna/US6035 COMB.seq.*
91: /cgn2_6/ptodata/1/pna/US6036 COMB.seq.*
92: /cgn2_6/ptodata/1/pna/US6037 COMB.seq.*
93: /cgn2_6/ptodata/1/pna/US6038 COMB.seq.*
94: /cgn2_6/ptodata/1/pna/US6039 COMB.seq.*
95: /cgn2_6/ptodata/1/pna/US6040 COMB.seq.*
96: /cgn2_6/ptodata/1/pna/US6041 COMB.seq.*
97: /cgn2_6/ptodata/1/pna/US6042 COMB.seq.*
98: /cgn2_6/ptodata/1/pna/US6043 COMB.seq.*
99: /cgn2_6/ptodata/1/pna/US6044 COMB.seq.*
100: /cgn2_6/ptodata/1/pna/US6045 COMB.seq.*
101: /cgn2_6/ptodata/1/pna/US6046 COMB.seq.*
102: /cgn2_6/ptodata/1/pna/US6047 COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2709	100.0	2709	33	US-09-774-490-1
2	2517	92.9	2848	48	US-10-247-671-83
3	2517	92.9	2848	71	US-60-172-373-8935
4	2517	92.9	2848	87	US-60-323-784-83

Sequence 1, Appl
Sequence 83, Appl
Sequence 8935, Ap
Sequence 83, Appl

```
5 2508.4 92.6 2530 1 PCT-US02-07826-283
6 2508.4 92.6 2530 15 US-09-053-375B-1350
7 2508.4 92.6 2530 45 US-10-097-340-283
8 2508.4 92.6 2530 49 US-10-262-538-9
9 2508.4 92.6 2601 9 US-08-488-935-1
10 2508.4 92.6 2601 9 US-08-489-057-1
11 2508.4 92.6 2601 45 US-10-067-632-53
12 2502.4 92.4 2857 47 US-10-170-235-19890
13 2050 75.7 2913 46 US-10-144-771-9178
14 2050 75.7 2913 91 US-60-360-207-9178
15 1544.8 57.0 2256 28 US-09-649-164-8352
16 1544.8 57.0 2256 28 US-09-652-816-7308
17 1544.8 57.0 2256 28 US-09-652-918-7516
18 817 30.2 944 32 US-09-724-676-18078
19 817 30.2 944 32 US-09-724-676A-18078
20 718.6 26.5 2331 25 US-09-582-419A-2
21 718.6 26.5 2331 52 US-10-404-438-2
22 718.6 26.5 3871 40 US-09-946-374-309
23 718.6 26.5 3871 44 US-10-006-041A-309
24 718.6 26.5 3871 44 US-10-006-063A-309
25 718.6 26.5 3871 44 US-10-006-116A-309
26 718.6 26.5 3871 44 US-10-006-117A-309
27 718.6 26.5 3871 44 US-10-006-130A-309
28 718.6 26.5 3871 44 US-10-006-172A-309
29 718.6 26.5 3871 44 US-10-006-485A-309
30 718.6 26.5 3871 44 US-10-006-746A-309
31 718.6 26.5 3871 44 US-10-006-768A-309
32 718.6 26.5 3871 44 US-10-006-818A-309
33 718.6 26.5 3871 44 US-10-006-856A-309
34 718.6 26.5 3871 44 US-10-007-194A-309
35 718.6 26.5 3871 44 US-10-007-236A-309
36 718.6 26.5 3871 44 US-10-011-671A-309
37 718.6 26.5 3871 44 US-10-011-692A-309
38 718.6 26.5 3871 44 US-10-011-795A-309
39 718.6 26.5 3871 44 US-10-011-795B-309
40 718.6 26.5 3871 44 US-10-011-833A-309
41 718.6 26.5 3871 44 US-10-012-064A-309
42 718.6 26.5 3871 44 US-10-012-101B-309
43 718.6 26.5 3871 44 US-10-012-121A-309
44 718.6 26.5 3871 44 US-10-012-137A-309
45 718.6 26.5 3871 44 US-10-012-149A-309
```

ALIGNMENTS

```
RESULT 1
US-09-774-490-1
; Sequence 1, Application US/09774490
; GENERAL INFORMATION:
; APPLICANT: Jin, Shengfang
; TITLE OF INVENTION: RESISTANCE SEQUENCES AND USES THEREOF
; FILE REFERENCE: 07334-138001
; CURRENT APPLICATION NUMBER: US/09/774,490
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: US 60/179,191
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2709
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-774-490-1
```

```
Query Match 100.0%; Score 2709; DB 33; Length 2709;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 AATCTTTTATCGATGTTAAACAGCTTGTAGTATCGATCCACGTCGAGGGTGTGC 60
Db 1 AATCTTTTATCGATGTTAAACAGCTTGTAGTATCGATCCACGTCGAGGGTGTGC 60
```

```
Qy 61 ACCCAGCGTCCGGGAGTAGTTGAGCTCGCTGCTTCTCCCATTTGTCCAGCCAGTCTATTT 120
Db 61 ACCCAGCGTCCGGGAGTAGTTGAGCTCGCTGCTTCTCCCATTTGTCCAGCCAGTCTATTT 120
Qy 121 CCAGATTGTTTGAATCTTCTGCGCGGCAATACAGGAAGGAAGACTAAAGCAGCAAGG 180
Db 121 CCAGATTGTTTGAATCTTCTGCGCGGCAATACAGGAAGGAAGACTAAAGCAGCAAGG 180
Qy 181 GACCTACAGCGTCTGCAGCATGGGCTGTTAATAGGATGTTCTCTCTTTCTTGGGGAGT 240
Db 181 GACCTACAGCGTCTGCAGCATGGGCTGTTAATAGGATGTTCTCTCTTTCTTGGGGAGT 240
Qy 241 ATTACTACAGCAAGAGCAAACTATCAGATGGGAAGCAAACTATGTCGCAAGGCTGAAAT 300
Db 241 ATTACTACAGCAAGAGCAAACTATCAGATGGGAAGCAAACTATGTCGCAAGGCTGAAAT 300
Qy 301 ATCTTACAAAAGAAATGTTGGAATCCAAATGTGTATCACTTTCAATGGCTGGCCAAACAG 360
Db 301 ATCTTACAAAAGAAATGTTGGAATCCAAATGTGTATCACTTTCAATGGCTGGCCAAACAG 360
Qy 361 CTCCAGTTTATCATACCTTCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAA 420
Db 361 CTCCAGTTTATCATACCTTCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAA 420
Qy 421 GGATCACATATTTTCAATTCGACCTGTTTAAATATCAAGGATTTTCAAAAAGATTTGTGGCC 480
Db 421 GGATCACATATTTTCAATTCGACCTGTTTAAATATCAAGGATTTTCAAAAAGATTTGTGGCC 480
Qy 481 AGTATTTTACACAGAGAGATGAATGCAATGGGCTGGAAAGACATCTCTGAAAGATG 540
Db 481 AGTATTTTACACAGAGAGATGAATGCAATGGGCTGGAAAGACATCTCTGAAAGATG 540
Qy 541 TGCTAAATTTTCAATCAAGGTACTTAAAGGCATATAAATCAGACTCACTTGTAGCCTGTGGAC 600
Db 541 TGCTAAATTTTCAATCAAGGTACTTAAAGGCATATAAATCAGACTCACTTGTAGCCTGTGGAC 600
Qy 601 GGGGGCTTTTTCATCCAAATTTGACCTATCAATTTGAAATTTGACATCATCTCTGAGGACAAT 660
Db 601 GGGGGCTTTTTCATCCAAATTTGACCTATCAATTTGAAATTTGACATCATCTCTGAGGACAAT 660
Qy 661 TTTTAAGCTGAGAACTCACAATTTTGAACCGGCGTGGGAAGAGTCCATATGACCTTAA 720
Db 661 TTTTAAGCTGAGAACTCACAATTTTGAACCGGCGTGGGAAGAGTCCATATGACCTTAA 720
Qy 721 GCTGCTGACAGCTCCCTTTTAAATAGATGGAGAAATTAATCTGCACTGAGCTGAGCTGATTT 780
Db 721 GCTGCTGACAGCTCCCTTTTAAATAGATGGAGAAATTAATCTGCACTGAGCTGAGCTGATTT 780
Qy 781 TATGGGCGAGACTTTTGTCTATCTTCCGAACTCTTTGGGCACCAACCACCAATCAGGACAGA 840
Db 781 TATGGGCGAGACTTTTGTCTATCTTCCGAACTCTTTGGGCACCAACCACCAATCAGGACAGA 840
Qy 841 GCAGCATGATTTCCAGGTGGCTCAATGATCCAAAGTTCAATAGTCCCACTCATCTCAGA 900
Db 841 GCAGCATGATTTCCAGGTGGCTCAATGATCCAAAGTTCAATAGTCCCACTCATCTCAGA 900
Qy 901 GAGTGACAACTCCTGAAGATGACAAAGTATATCTTTTCCGTAAGAAATGCAATAGATGG 960
Db 901 GAGTGACAACTCCTGAAGATGACAAAGTATATCTTTTCCGTAAGAAATGCAATAGATGG 960
Qy 961 AGAACACTCTCGAAAAAGCTACTCAGCTAGATAGGTCAGATATGCAAGAATGACTTTGG 1020
Db 961 AGAACACTCTCGAAAAAGCTACTCAGCTAGATAGGTCAGATATGCAAGAATGACTTTGG 1020
Qy 1021 AGGGCAGAAAGTCTGGTGAATAAATGGAACAACTTCTTCAAGCTCTGCTGATTTGCTC 1080
Db 1021 AGGGCAGAAAGTCTGGTGAATAAATGGAACAACTTCTTCAAGCTCTGCTGATTTGCTC 1080
Qy 1081 AGTCCAGGTCCTCAATGCACTTATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1140
Db 1081 AGTCCAGGTCCTCAATGCACTTATTTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1140
Qy 1141 GAACTTTTAAAGATCTTAAAAATCCAGTTTGTATATGGAGTGTGTTTACGACTTCCAGTAACAT 1200
```


1569 TATTATCGGAAAGCGTGTGCTGAGTGTGCTCGCCCGAGACCTTACTGTGCTGGGA 1628
1813 TGGTCTGCTATGTTCTCGCTATTTCCCACTCAAGAGACGACAGACGACGATAT 1872
1629 TGGTCTGCTATGTTCTCGCTATTTCCCACTCAAGAGACGACAGACGACGATAT 1688
1873 AAGAAATGAGAGACCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1932
1689 AAGAAATGAGAGACCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1748
1933 CAGCCCTGAGAGAGAGATCATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1992
1749 CAGCCCTGAGAGAGAGATCATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1808
1993 TCGAAGTCCGAGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2052
1809 TCGAAGTCCGAGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1868
2053 AAGAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGGTAG 2112
1869 AAGAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGGTAG 1928
2113 TCTAACACAGAGAGATTCAGGCAATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2172
1929 TCTAACACAGAGAGATTCAGGCAATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1988
2173 AACTCTTCTTAAGTGAACCTGGAAGTCAATGACAGAGAGATTCAGGCAATTCATCA 2232
1989 AACTCTTCTTAAGTGAACCTGGAAGTCAATGACAGAGAGATTCAGGCAATTCATCA 2048
2233 TAAAGATGATGAGAGTGGCTCTAAGACCAAGAAATGCTCAATGATGACACCTAG 2292
2049 TAAAGATGATGAGAGTGGCTCTAAGACCAAGAAATGCTCAATGATGACACCTAG 2108
2293 CCAGAGAGTCTGGTACAGAGACTTCATGAGCTGCTCAACCCCAATCTCAACAGCAT 2352
2109 CCAGAGAGTCTGGTACAGAGACTTCATGAGCTGCTCAACCCCAATCTCAACAGCAT 2168
2353 GATGAGTCTGTGAACAGTGTGGAAGAGGACCGAAACCAACCTGCGCAAGGCGAG 2412
2169 GATGAGTCTGTGAACAGTGTGGAAGAGGACCGAAACCAACCTGCGCAAGGCGAG 2228
2413 ACATACCCAGGAGACAGTAAACAAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAA 2472
2229 ACATACCCAGGAGACAGTAAACAAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAA 2288
2473 CAGGAGACCCAGATTTGAGAGGACCCAGAGGCTGCTGAGCTGCTGCTGCTGCTGCTGCT 2532
2289 CAGGAGACCCAGATTTGAGAGGACCCAGAGGCTGCTGAGCTGCTGCTGCTGCTGCTGCT 2348
2533 AACCTCAACAGTAGAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2592
2349 AACCTCAACAGTAGAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2408
2593 GAACTTTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2652
2409 GAACTTTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2468
2653 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2709
2469 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2525

RESULT 4

US-60-323-784-83

; Sequence 83, Application US/60323784
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS

; FILE REFERENCE: PA-0050 P
; CURRENT APPLICATION NUMBER: US/60/323,784
; CURRENT FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 83
; LENGTH: 2848
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc. feature
; OTHER INFORMATION: Incyte ID No: 1930967CB1
US-60-323-784-83

Query Match 92.9%; Score 2517; DB 87; Length 2848;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2517; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 CTGAGCAGATGGGCTGGTTAACTAGGATTTGCTGCTTTTCTGGGAGATTAATCTTACAGC 252
DB 9 CTGAGCAGATGGGCTGGTTAACTAGGATTTGCTGCTTTTCTGGGAGATTAATCTTACAGC 68
QY 253 AAGAGCAAACTATCAGAAATGGGAAGAACAAATGTGCAAGGCTGAAATTAATCTTACAGAG 312
DB 69 AAGAGCAAACTATCAGAAATGGGAAGAACAAATGTGCAAGGCTGAAATTAATCTTACAGAG 128
QY 313 AATGTTGGAATCCCAACATGTGATCCTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCA 372
DB 129 AATGTTGGAATCCCAACATGTGATCCTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCA 188
QY 373 TACCTTCCTTTTGGATGAGGAAACGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATATT 432
DB 189 TACCTTCCTTTTGGATGAGGAAACGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATATT 248
QY 433 TTCATTGCACTTGGTTAAATATCAAGATTTTCAAAAGATTTGTTGGCCAGATTTCTTACAC 492
DB 249 TTCATTGCACTTGGTTAAATATCAAGATTTTCAAAAGATTTGTTGGCCAGATTTCTTACAC 308
QY 493 CAGAGAGATCAATGCAAGTGGCTGGAAAGACATCTGAAAGAAATGTGCTAAATTTTCAAT 552
DB 309 CAGAGAGATCAATGCAAGTGGCTGGAAAGACATCTGAAAGAAATGTGCTAAATTTTCAAT 368
QY 553 CAAGTACTTAAGGCATATAATCAGACTCAGCTGTGTACGCTGTGTGAAACGGGGCTTTTCA 612
DB 369 CAAGTACTTAAGGCATATAATCAGACTCAGCTGTGTACGCTGTGTGAAACGGGGCTTTTCA 428
QY 613 TCCAAATTTGCACTTACATTTGAAATTTGACATCATCTGAGGACATATTTTAAAGCTGGA 672
DB 429 TCCAAATTTGCACTTACATTTGAAATTTGACATCATCTGAGGACATATTTTAAAGCTGGA 488
QY 673 GAACTCAATTTTGAACACGGCTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGC 732
DB 489 GAACTCAATTTTGAACACGGCTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGC 548
QY 733 ATCCCTTTTAATAGATGGAGAAATTAATCTGGAATCTGAGCTGATTTTATGGGGCGAGA 792
DB 549 ATCCCTTTTAATAGATGGAGAAATTAATCTGGAATCTGAGCTGATTTTATGGGGCGAGA 608
QY 793 CTTTGTCTATCTTCGAACTCTTGGGACCAACCAATCAGGACAGAGCAGCATGATTC 852
DB 609 CTTTGTCTATCTTCGAACTCTTGGGACCAACCAATCAGGACAGAGCAGCATGATTC 668
QY 853 CAGTGGCTCAATGATCCAAAGTTTCAATGAGTGGCCACCTCATCTCAGAGAGTGACATCC 912
DB 669 CAGTGGCTCAATGATCCAAAGTTTCAATGAGTGGCCACCTCATCTCAGAGAGTGACATCC 728
QY 913 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGAGAGAACTCTGG 972
DB 729 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGAGAGAACTCTGG 788
QY 973 AAAAGCTACTCAGCTAGAAATAGTTCAGATATGCAAGATGACATTTTGGAGGCGACAGAG 1032
DB 789 AAAAGCTACTCAGCTAGAAATAGTTCAGATATGCAAGATGACATTTTGGAGGCGACAGAG 848

Query Match	92.6%;	Score 2508.4;	DB 1;	Length 2530;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches 2509;	Conservative	0;	Mismatches	1;	Indels 0; Gaps 0;
QY	193	CTGCAGCATGGCGCTGGTTAACTAGGATGTCTGTCCTTTCTGCGGAGGTATTACTTACAGC	252		
DB	9	CTGCAGCATGGCGCTGGTTAACTAGGATGTCTGTCCTTTCTGCGGAGGTATTACTTACAGC	68		
QY	253	AAGAGCAAACTATCAGAAATGGGAAGAAACAATGTGCCAAGGCTGAAATTTATCTTACAAAGA	312		
DB	69	AAGAGCAAACTATCAGAAATGGGAAGAAACAATGTGCCAAGGCTGAAATTTATCTTACAAAGA	128		
QY	313	AATGTTGGAATCCAAACAATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCA	372		
DB	129	AATGTTGGAATCCAAACAATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCA	188		
QY	373	TACCTTCCTTTTGGATGAGGAACGAGTAGGCTGTATGTTGGAGCAAGGATCACATATT	432		
DB	189	TACCTTCCTTTTGGATGAGGAACGAGTAGGCTGTATGTTGGAGCAAGGATCACATATT	248		
QY	433	TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGTGGCCAGGTATCTTACAC	492		
DB	249	TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGTGGCCAGGTATCTTACAC	308		
QY	493	CAGAGAGATGAATGCAAGTGGGCTGGAAAAGACATCTCTGAAAAGAAATGTGCTAATTTTCA	552		
DB	309	CAGAGAGATGAATGCAAGTGGGCTGGAAAAGACATCTCTGAAAAGAAATGTGCTAATTTTCA	368		
QY	553	CAAGGTACTTAAAGCATATATCAGACTCATCTGTGACGCTGTGGAAACGGGGCTTTTCA	612		
DB	369	CAAGGTACTTAAAGCATATATATCAGACTCATCTGTGACGCTGTGGAAACGGGGCTTTTCA	428		
QY	613	TCCAAATTTGCACCTACATTTGAAATTTGGACATCATCTCTGAGGACAATATTTTAAAGCTGGA	672		
DB	429	TCCAAATTTGCACCTACATTTGAAATTTGGACATCATCTCTGAGGACAATATTTTAAAGCTGGA	488		
QY	673	GAACTCACATTTTGAACCGCGCTGGGAAGATTCATATCAACCTAAGCTGCTGCAGC	732		
DB	489	GAACTCACATTTTGAACCGCGCTGGGAAGATTCATATCAACCTAAGCTGCTGCAGC	548		
QY	733	ATCCCTTTTAAATGAGAAATTAATCTCTGGAACCTGCAGCTGATTTTATGGGGGAG	792		
DB	549	ATCCCTTTTAAATGAGAAATTAATCTCTGGAACCTGCAGCTGATTTTATGGGGGAG	608		
QY	793	CTTTGCTATCTTCGAACTCTTTGGGCACACACCCAAATCAGGACAGACAGATGATTC	852		
DB	609	CTTTGCTATCTTCGAACTCTTTGGGCACACACCCAAATCAGGACAGACAGATGATTC	668		
QY	853	CAGGTGGCTCAATGATCCAAAGTTCATTTAGTGGCCACCTCATCTCAGAGAGTGACAATCC	912		
DB	669	CAGGTGGCTCAATGATCCAAAGTTCATTTAGTGGCCACCTCATCTCAGAGAGTGACAATCC	728		
QY	913	TGAAGATGACAAAGTATATCTTTCTCCGTTGAAAATGCAATAGATGGAGAACCTCTGG	972		
DB	729	TGAAGATGACAAAGTATATCTTTCTCCGTTGAAAATGCAATAGATGGAGAACCTCTGG	788		
QY	973	AAAAGCTACTCAGCTAGAAATAGGTTCAGATATGCAAGAAATGCACTTTGGAGGACACAAAG	1032		
DB	789	AAAAGCTACTCAGCTAGAAATAGGTTCAGATATGCAAGAAATGCACTTTGGAGGACACAAAG	848		
QY	1033	TCGTGTGAATAATGGACAAACATTCCTCAAAAGCTCGTCTGATTTGCTCAGTGGCAGGTCC	1092		
DB	849	TCGTGTGAATAATGGACAAACATTCCTCAAAAGCTCGTCTGATTTGCTCAGTGGCAGGTCC	908		
QY	1093	AAATGGCATTTGACATCTATTTTGTATGAACTGAGATGTATTCCTTAATGAACTTTAAAGA	1152		
DB	909	AAATGGCATTTGACATCTATTTTGTATGAACTGAGATGTATTCCTTAATGAACTTTAAAGA	968		
QY	1153	TCTTAAATCCAGTTGTATATGAGTCTGTTTACGACTTCCAGTAAACATTTTCAAGGATC	1212		
DB	969	TCTTAAATCCAGTTGTATATGAGTCTGTTTACGACTTCCAGTAAACATTTTCAAGGATC	1028		
QY	1213	AGCCGTGTGTATGTATGATCATGATGTGTGAGAAGGGTGTTCCTGCTCATATGCCCA	1272		

1449 AGAAATGACAGATTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTTCCACTAAGCA 1508
1693 GCAACAACTATATATTTGGTTCAACGGCTGGGTTGCCAGCTCCCTTTACACGGGTGCA 1752
1509 GCAACAACTATATATTTGGTTCAACGGCTGGGTTGCCAGCTCCCTTTACACGGGTGCA 1568
1753 TATTTACGGGAAGCGTGTGAGTGTTCCTCGCCGAGACCCCTTACTGTCTTGGGA 1812
1569 TATTTACGGGAAGCGTGTGAGTGTTCCTCGCCGAGACCCCTTACTGTCTTGGGA 1628
1813 TGTTCTCGATGTTCTCGCTATTTTCCCACTGCAAGAGACGCAAGACGCAAGATAT 1872
1629 TGTTCTCGATGTTCTCGCTATTTTCCCACTGCAAGAGACGCAAGACGCAAGATAT 1688
1873 AAGAAATGAGAGACCCACTGACTCAGCTGTTTCCAGCTTACACCATGATATACCATGGCA 1932
1689 AAGAAATGAGAGACCCACTGACTCAGCTGTTTCCAGCTTACACCATGATATACCATGGCA 1748
1933 CAGCCCTGAAGAGAGATCATCTATGTTGTAGAGATAGTACACATTTTTCGAATGCAG 1992
1749 CAGCCCTGAAGAGAGATCATCTATGTTGTAGAGATAGTACACATTTTTCGAATGCAG 1808
1993 TCCGAAGTCCGAGAGAGCGCTGCTCTATTTGGCAATTTCCAGAGCGCAAAATGAAGAGCGAA 2052
1809 TCCGAAGTCCGAGAGAGCGCTGCTCTATTTGGCAATTTCCAGAGCGCAAAATGAAGAGCGAA 1868
2053 AAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGTAG 2112
1869 AAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGTAG 1928
2113 TCTACAAACAGAGATTTCAAGCAATTTACCTCTGCCATGGGTGGCAATGGTTTATACA 2172
1929 TCTACAAACAGAGATTTCAAGCAATTTACCTCTGCCATGGGTGGCAATGGTTTATACA 1988
2173 AACTCTTCTTAAGGTAAACCTCGAAGTCAATGACACAGAGCATTTGGGAACATCTTTCA 2232
1989 AACTCTTCTTAAGGTAAACCTCGAAGTCAATGACACAGAGCATTTGGGAACATCTTTCA 2048
2233 TAAAGTATGATGGAGATGGCTCTTAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2292
2049 TAAAGTATGATGGAGATGGCTCTTAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2108
2293 CAGAGAGTCTGTGACAGAGATTTCAATGAGCTTCAATCAACACCCCAATCTCAACAGAT 2352
2109 CAGAGAGTCTGTGACAGAGATTTCAATGAGCTTCAATCAACACCCCAATCTCAACAGAT 2168
2353 GGATGAGTCTGTGAAACAAAGTTTGGAAAAGGGACCGAAACCAACGTCGCAAGGCCAGG 2412
2169 GGATGAGTCTGTGAAACAAAGTTTGGAAAAGGGACCGAAACCAACGTCGCAAGGCCAGG 2228
2413 ACATACCCAGGAGACAGTAAACAAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAA 2472
2229 ACATACCCAGGAGACAGTAAACAAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAA 2288
2473 CAGGAGACCCAGAAATTTGAGAGGACCCAGGAGTGTCTGAGCTGCAATACCTCTAGA 2532
2289 CAGGAGACCCAGAAATTTGAGAGGACCCAGGAGTGTCTGAGCTGCAATACCTCTAGA 2348
2533 AACCTCAAAACAGTAGAAACTTTCCTCTAGCAATTAACCTGGAAGAAACAAATGCAATATACAT 2592
2349 AACCTCAAAACAGTAGAAACTTTCCTCTAGCAATTAACCTGGAAGAAACAAATGCAATATACAT 2408
2593 GAACCTTTTTCAGGCAATATGAGATGTTTCAATGTTGGGAAATTCAGCTGAGTTCCA 2652
2409 GAACCTTTTTCAGGCAATATGAGATGTTTCAATGTTGGGAAATTCAGCTGAGTTCCA 2468
2653 CCAATTAATAATTAATCCATGAGTAACTTTCTCTAATAGGCTTTTTTTC 2702
2469 CCAATTAATAATTAATCCATGAGTAACTTTCTCTAATAGGCTTTTTTTC 2518

RESULT 7

US-10-097-340-283
; Sequence 283, Application US/10097340
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAWATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHWANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 283
; LENGTH: 2530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-097-340-283

Query Match 92.6%; Score 2508.4; DB 45; Length 2530;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 193 CTGCAGCATGGCTGTTAACTAGGATTTCTGTTCTTTCTGGGGAGTATTACTTTACAGC 252
Db 9 CTGCAGCATGGCTGTTAACTAGGATTTCTGTTCTTTCTGGGGAGTATTACTTTACAGC 68
Qy 253 AAGAGCAAACTATCAGAAATGGAAGCAAAATGTGCAAGGCTGAAATTTATCTTACAAAGA 312
Db 69 AAGAGCAAACTATCAGAAATGGAAGCAAAATGTGCAAGGCTGAAATTTATCTTACAAAGA 128
Qy 313 AATGTTGGAATCCCAACAAATGTGATCACTTTCAATGGCTTGGCCCAACAGCTCCAGTTATCA 372
Db 129 AATGTTGGAATCCCAACAAATGTGATCACTTTCAATGGCTTGGCCCAACAGCTCCAGTTATCA 188
Qy 373 TACCTCTCTTTGGATGAGGAACCGAGTAGCTGTATGTTGAGCAAAAGGATCACATATT 432
Db 189 TACCTCTCTTTGGATGAGGAACCGAGTAGCTGTATGTTGAGCAAAAGGATCACATATT 248
Qy 433 TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTTTGTGGCCAGATCTTTACAC 492
Db 249 TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTTTGTGGCCAGATCTTTACAC 308
Qy 493 CAGAAGAGATGAATGCAAGTGGGCTGGAAAGACATCTCTGAAAGATGTGCTTAATTTAT 552
Db 309 CAGAAGAGATGAATGCAAGTGGGCTGGAAAGACATCTCTGAAAGATGTGCTTAATTTAT 368

QY 553 CAAAGTACTTAAGGCATATAATCAGACTCATTGTAGCGCTGTGGAAACGGGGCGCTTTTCA 612
DB 369 CAAAGTACTTAAGGCATATAATCAGACTCATTGTAGCGCTGTGGAAACGGGGCGCTTTTCA 428
QY 613 TCCAAATTTGCACCTACATTGAAATTTGGACATCATCTGAGGACAATATTTTAAAGCTGGA 672
DB 429 TCCAAATTTGCACCTACATTGAAATTTGGACATCATCTGAGGACAATATTTTAAAGCTGGA 488
QY 673 GAACTCACATTTTGAACCGCGGTGGGAAGTCCATATGACCCCTAAAGCTGCTGACAGC 732
DB 489 GAACTCACATTTTGAACCGCGGTGGGAAGTCCATATGACCCCTAAAGCTGCTGACAGC 548
QY 733 ATCCCTTTTATAGATGAGAAATATATCTCTGGAATCTGAGCTGATTTTATGGGGCGGAGA 792
DB 549 ATCCCTTTTATAGATGAGAAATATATCTCTGGAATCTGAGCTGATTTTATGGGGCGGAGA 608
QY 793 CTTCCTGCTATCTCCGAACTCTTTGGGACACCCACCCCAATCAGGACAGAGCGACATGATTC 852
DB 609 CTTCCTGCTATCTCCGAACTCTTTGGGACACCCACCCCAATCAGGACAGAGCGACATGATTC 668
QY 853 CAGGTGGCTCAATGATCCAAAGTTCATTTAGTCCCACTCTCATCTCAGAGAGTGACAAATCC 912
DB 669 CAGGTGGCTCAATGATCCAAAGTTCATTTAGTCCCACTCTCATCTCAGAGAGTGACAAATCC 728
QY 913 TGAAGATGACAAAGTATATCTTTTCTTCGGTGAATAATGCAATAGATGAGAAACACTCTGG 972
DB 729 TGAAGATGACAAAGTATATCTTTTCTTCGGTGAATAATGCAATAGATGAGAAACACTCTGG 788
QY 973 AAAAGCTACTCAGCTAGAAATGAGTGCAGATGCAAGATGATGATTTGGAGGGGACAGAA 1032
DB 789 AAAAGCTACTCAGCTAGAAATGAGTGCAGATGCAAGATGATGATTTGGAGGGGACAGAA 848
QY 1033 TCTGTGTGAATAAATGGACAACTTCCTCAAGCTCGTCTGATTTGCTCAGTGCACAGTCC 1092
DB 849 TCTGTGTGAATAAATGGACAACTTCCTCAAGCTCGTCTGATTTGCTCAGTGCACAGTCC 908
QY 1093 AAATGGCAATGACACTCACTTTTGTATGAACTGACAGGATGATTTCTTAATGAACTTTAAAGA 1152
DB 909 AAATGGCAATGACACTCACTTTTGTATGAACTGACAGGATGATTTCTTAATGAACTTTAAAGA 968
QY 1153 TCCTAAAATCCAGTTGTATATGAGTGTGTACGACTTCCAGTACATTTTCAAGGGATC 1212
DB 969 TCCTAAAATCCAGTTGTATATGAGTGTGTACGACTTCCAGTACATTTTCAAGGGATC 1028
QY 1213 AGCCGTGTGTATATGATGATGAGTGTGTAGAGGGTGTTCCTGTGCTCATATGCCCA 1272
DB 1029 AGCCGTGTGTATATGATGATGAGTGTGTAGAGGGTGTTCCTGTGCTCATATGCCCA 1088
QY 1273 CAGGATGGACCCCACTATCAATGGTGCCTTTATCAAGGAAAGTTCCTCATGCCGCC 1332
DB 1089 CAGGATGGACCCCACTATCAATGGTGCCTTTATCAAGGAAAGTTCCTCATGCCGCC 1148
QY 1333 AGGAACTTGTCCAGCAAAACATTTGGTGGTGTGTGACTCTCAAGAGGACCTTCCTGATGA 1392
DB 1149 AGGAACTTGTCCAGCAAAACATTTGGTGGTGTGTGACTCTCAAGAGGACCTTCCTGATGA 1208
QY 1393 TGTATAACCTTTGCAAGAGTCACTCCAGGCAATGACAAATCCAGTGTTCCTATGAACAA 1452
DB 1209 TGTATAACCTTTGCAAGAGTCACTCCAGGCAATGACAAATCCAGTGTTCCTATGAACAA 1268
QY 1453 TCGCCCAATAGTGTATCAAAACGGATGTAAATTTATCAATTTTACAAAATGTCGTAGACCG 1512
DB 1269 TCGCCCAATAGTGTATCAAAACGGATGTAAATTTATCAATTTTACAAAATGTCGTAGACCG 1328
QY 1513 AGTGAATGCAAGATGACAGATGATGATTTATGTTTATCGGAACAGATGTTGGACCGT 1572
DB 1329 AGTGAATGCAAGATGACAGATGATGATTTATGTTTATCGGAACAGATGTTGGACCGT 1388
QY 1573 TCTTAAAGTGTTCCTTCAAGGACCTTGGTATGATTTTGAAGAGGTTCTGCTGGA 1632
DB 1389 TCTTAAAGTGTTCCTTCAAGGACCTTGGTATGATTTTGAAGAGGTTCTGCTGGA 1448

QY 1633 AGAAATGACAGTTCCTCGGAAACCGACTGTCTATTTTACGAATGGAGCTTTTCCACTAAGCA 1692
DB 1449 AGAAATGACAGTTCCTCGGAAACCGACTGTCTATTTTACGAATGGAGCTTTTCCACTAAGCA 1508
QY 1693 GCAACAACTATATTTGGTTTCAACGGCTGGGGTTGCCAGCTCCCTTTTACACCGGTGGA 1752
DB 1509 GCAACAACTATATTTGGTTTCAACGGCTGGGGTTGCCAGCTCCCTTTTACACCGGTGGA 1568
QY 1753 TATTTACGGGAAAGCGTGTGCTGAGTGTTCCTCGCCGAGACCCCTTACTGTCTTGGGA 1812
DB 1569 TATTTACGGGAAAGCGTGTGCTGAGTGTTCCTCGCCGAGACCCCTTACTGTCTTGGGA 1628
QY 1813 TGGTTCTGCTATGTTCTCGCTATTTTCCCACTGTGAAAAGAGACGCAACAGCAAGATAT 1872
DB 1629 TGGTTCTGCTATGTTCTCGCTATTTTCCCACTGTGAAAAGAGACGCAACAGCAAGATAT 1688
QY 1873 AAGAAATGGACACCCACTGACTCTGTTTCCAGACTTACACCATGATATACCATGGCCA 1932
DB 1689 AAGAAATGGACACCCACTGACTCTGTTTCCAGACTTACCATGATATACCATGGCCA 1748
QY 1933 CAGCCCTGAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGCAATTTTGGAAATGCAG 1992
DB 1749 CAGCCCTGAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGCAATTTTGGAAATGCAG 1808
QY 1993 TCCGAAGTCCGAGAGAGCGCTGTTTATTTGGCAATTTCCAGAGGGGAAATGAAGAGCGAAA 2052
DB 1809 TCCGAAGTCCGAGAGAGCGCTGTTTATTTGGCAATTTCCAGAGGGGAAATGAAGAGCGAAA 1868
QY 2053 AGAAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCCCTTCTGCTACGTTAG 2112
DB 1869 AGAAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCCCTTCTGCTACGTTAG 1928
QY 2113 TCTACACAGAGAGGATTCAGGCAATTTACTCTGCCATCGGTGGAAACATGGTTCATACA 2172
DB 1929 TCTACACAGAGAGGATTCAGGCAATTTACTCTGCCATCGGTGGAAACATGGTTCATACA 1988
QY 2173 AACTCTTCTTAAAGTAAACCTCGGAAAGTCAATTTGACAGAGCATTTGGAAAGAACTTCTTCA 2232
DB 1989 AACTCTTCTTAAAGTAAACCTCGGAAAGTCAATTTGACAGAGCATTTGGAAAGAACTTCTTCA 2048
QY 2233 TAAAGATCAGATCGAGATGGCTCTTAAGACCAAGAAATGTCCAATAGTACGACACCTAG 2292
DB 2049 TAAAGATCAGATCGAGATGGCTCTTAAGACCAAGAAATGTCCAATAGTACGACACCTAG 2108
QY 2293 CCAGAGGTCTGGTACAGAGACTTTCATGCGACTCATCAACCAACCCCAATCTCAACAGCAT 2352
DB 2109 CCAGAGGTCTGGTACAGAGACTTTCATGCGACTCATCAACCAACCCCAATCTCAACAGCAT 2168
QY 2353 GGATGAGTTCGTGAAACAAAGTTTGGAAAGGGACCGGAAACAAACGTGGCAAGGCCAGG 2412
DB 2169 GGATGAGTTCGTGAAACAAAGTTTGGAAAGGGACCGGAAACAAACGTGGCAAGGCCAGG 2228
QY 2413 ACATACCCAGGAAACAGTAACTTGAAGGACCTTACAGAAATTAAGAAAGGTAGAAA 2472
DB 2229 ACATACCCAGGAAACAGTAACTTGAAGGACCTTACAGAAATTAAGAAAGGTAGAAA 2288
QY 2473 CAGGAGGACCCCAAGTTTGGAGGGCACCCAGGAGTGTCTGAGCTGCATTACCTCTAGA 2532
DB 2289 CAGGAGGACCCCAAGTTTGGAGGGCACCCAGGAGTGTCTGAGCTGCATTACCTCTAGA 2348
QY 2533 AACCTCAAAACAGTAGAAACTTTCCTAGACAAATTAACCTGGAAAAACAAATGCAATATACAT 2592
DB 2349 AACCTCAAAACAGTAGAAACTTTCCTAGACAAATTAACCTGGAAAAACAAATGCAATATACAT 2408
QY 2593 GAACCTTTTTCATGGCAATTTATGGGATGTTTACAACTGGTGGGAAATTCAGCTGAGTTCCA 2652
DB 2409 GAACCTTTTTCATGGCAATTTATGGGATGTTTACAACTGGTGGGAAATTCAGCTGAGTTCCA 2468
QY 2653 CCAATTAATAATTAATCCCATGAGTAACTTTCCTTAATAGGCTTTTTCCTCC 2702
DB 2469 CCAATTAATAATTAATCCCATGAGTAACTTTCCTTAATAGGCTTTTTCCTCC 2518

RESULT 8

US-10-262-538-9
; Sequence 9, Application US/10262538
; GENERAL INFORMATION:
; APPLICANT: Alitalo et al
; TITLE OF INVENTION: NEUROFILIN/VEGF-C/VEGFR-3 MATERIALS AND METHODS
; FILE REFERENCE: 28967/37564
; CURRENT APPLICATION NUMBER: US/10/262,538
; CURRENT FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
; LENGTH: 2530
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (16)..(2331)
US-10-262-538-9

Query Match 92.6%; Score 2508.4; DB 49; Length 2530;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	193	CTGCAGCATGGGCTGTTAACTAGGATTCCTCTCTTTCTGGGAGATTAATCTTACAGC	252
DB	9	CTGCAGCATGGGCTGTTAACTAGGATTCCTCTCTTTCTGGGAGATTAATCTTACAGC	68
QY	253	AAGAGCAAACTATCAGAAATGGGAAGAACTATGCGCAAGCTGAAATTAATCTCAAAAGA	312
DB	69	AAGAGCAAACTATCAGAAATGGGAAGAACTATGCGCAAGCTGAAATTAATCTCAAAAGA	128
QY	313	AATGTTGGAAATCCAAATGTCATCTTCAATGCTTGGCCACACAGCTCCAGTTATCA	372
DB	129	AATGTTGGAAATCCAAATGTCATCTTCAATGCTTGGCCACACAGCTCCAGTTATCA	188
QY	373	TACCTTCTCTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAAAGGATCAATAT	432
DB	189	TACCTTCTCTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAAAGGATCAATAT	248
QY	433	TTCAATTCGACCTGGTTAAATCAAGGATTTCAAAAGATTTGTTGGCCAGTATCTTACAC	492
DB	249	TTCAATTCGACCTGGTTAAATCAAGGATTTCAAAAGATTTGTTGGCCAGTATCTTACAC	308
QY	493	CAGAGAGATGAATCAAGTGGGCTGGAAAGACATCTCTGAAGAAATGCTGAATTTCTAT	552
DB	309	CAGAGAGATGAATCAAGTGGGCTGGAAAGACATCTCTGAAGAAATGCTGAATTTCTAT	368
QY	553	CAAGGTACTTTAAGGCATATAATCAGACTCACTTTGACGCTGTGGAAACGGGGCTTTTCA	612
DB	369	CAAGGTACTTTAAGGCATATAATCAGACTCACTTTGACGCTGTGGAAACGGGGCTTTTCA	428
QY	613	TCCAATTCGACCTAATTTGAATTTGGAATCATCTCTGAGGACATATTTTAAAGCTGGA	672
DB	429	TCCAATTCGACCTAATTTGAATTTGGAATCATCTCTGAGGACATATTTTAAAGCTGGA	488
QY	673	GAACTCACATTTTGAACCGGCGTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGC	732
DB	489	GAACTCACATTTTGAACCGGCGTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGC	548
QY	733	ATCCCTTTTAAATAGATGGAGAAATTAATCTCTGGAACCTGAGCTGATTTTATGGGCGAGA	792
DB	549	ATCCCTTTTAAATAGATGGAGAAATTAATCTCTGGAACCTGAGCTGATTTTATGGGCGAGA	608
QY	793	CTTTGCTATCTTCGAACTCTTTGGGCAACCAACCAATCAGACAGACAGCATGATTC	852
DB	609	CTTTGCTATCTTCGAACTCTTTGGGCAACCAACCAATCAGACAGACAGCATGATTC	668
QY	853	CAGGTGGCTCAATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGAATCC	912
DB	669	CAGGTGGCTCAATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGAATCC	728
QY	913	TGAAGATGACAAAGTATATCTTTTCTTCGTTGAAATGCAATAGATGGAGAACACTCTGG	972

DB	729	TGAAGATGACAAAGTATATCTTTCTTCGTTGAAATGCAATAGATGAGAACACTCTCG	788
QY	973	AAAAGCTACTCAGCTAGAAATAGGTACAGATATGCAAGAAATGACTTTGGAGGCGACAGAG	1032
DB	789	AAAAGCTACTCAGCTAGAAATAGGTACAGATATGCAAGAAATGACTTTGGAGGCGACAGAG	848
QY	1033	TCTGTGAAATAAATGGCAACAATCTCTCAAAAGCTGCTCTGATTTGCTCAGTCCAGGTC	1092
DB	849	TCTGTGAAATAAATGGCAACAATCTCTCAAAAGCTGCTCTGATTTGCTCAGTCCAGGTC	908
QY	1093	AAATGGCAATGCACACTCAATTTTGATGAATGCAGGATGATTTCTTAATGAATTTAAAGA	1152
DB	909	AAATGGCAATGCACACTCAATTTTGATGAATGCAGGATGATTTCTTAATGAATTTAAAGA	968
QY	1153	TCCTAAAATCCAGTTGATATGAGTGTTCAGACTTCCAGTACATTTTCAAGGGATC	1212
DB	969	TCCTAAAATCCAGTTGATATGAGTGTTCAGACTTCCAGTACATTTTCAAGGGATC	1028
QY	1213	AGCGTGTGATATGATAGCATGATGTCAGAAAGGCTTCTTGGTCCATATGCCCA	1272
DB	1029	AGCGTGTGATATGATAGCATGATGTCAGAAAGGCTTCTTGGTCCATATGCCCA	1088
QY	1273	CAGGATGGACCCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCAGGCC	1332
DB	1089	CAGGATGGACCCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCAGGCC	1148
QY	1333	AGGAACCTTGTCCAGCAAAACATTTGGTGGTTTGAACCTACAAAGGACCTTCTGATGA	1392
DB	1149	AGGAACCTTGTCCAGCAAAACATTTGGTGGTTTGAACCTACAAAGGACCTTCTGATGA	1208
QY	1393	TGTTATAACCTTTGCAAGAACTATCCAGCCATGTACAAATCCAGTGTTCCTATCAACAA	1452
DB	1209	TGTTATAACCTTTGCAAGAACTATCCAGCCATGTACAAATCCAGTGTTCCTATCAACAA	1268
QY	1453	TCGCCCAATAGTGTATCAAAACGGATGTAAATTTCAATTTACAAAATTTGCTGAGACG	1512
DB	1269	TCGCCCAATAGTGTATCAAAACGGATGTAAATTTCAATTTACAAAATTTGCTGAGACG	1328
QY	1513	AGTGATGCAAGATGACAGATATGATTTATGTTTATCGGAACAGATTTGGGACCGT	1572
DB	1329	AGTGATGCAAGATGACAGATATGATTTATGTTTATCGGAACAGATTTGGGACCGT	1388
QY	1573	TCCTAAAGTACTTTCAATTTCTAAGGACCTTGGTATGATTTAGAAAGGTTCTCTGGA	1632
DB	1389	TCCTAAAGTACTTTCAATTTCTAAGGACCTTGGTATGATTTAGAAAGGTTCTCTGGA	1448
QY	1633	AGAAATGACAGTCTTTTGGGAAACCGACTGCTATTTTACAGCAATGGAGCTTTCCACTA	1692
DB	1449	AGAAATGACAGTCTTTTGGGAAACCGACTGCTATTTTACAGCAATGGAGCTTTCCACTA	1508
QY	1693	GCAACAACTATATTTGTTTCAACGGCTGGGGTGGCCAGCTCCCTTTACACCGGTGGA	1752
DB	1509	GCAACAACTATATTTGTTTCAACGGCTGGGGTGGCCAGCTCCCTTTACACCGGTGGA	1568
QY	1753	TATTTACGGGAAAGCGTGTGCTGAGTGTGCTCGCCGAGACCTTACTGTCTTGGGA	1812
DB	1569	TATTTACGGGAAAGCGTGTGCTGAGTGTGCTCGCCGAGACCTTACTGTCTTGGGA	1628
QY	1813	TGGTTCTGCACTGTTCTCGCTATTTTCCCACTGCAAAAGAGACGCAACAGCAAGATAT	1872
DB	1629	TGGTTCTGCACTGTTCTCGCTATTTTCCCACTGCAAAAGAGACGCAACAGCAAGATAT	1688
QY	1873	AAGAAATGGAGACCCACTGACTCTGTTTACAGCTTACACCATGATATCACCATGGCCA	1932
DB	1689	AAGAAATGGAGACCCACTGACTCTGTTTACAGCTTACACCATGATATCACCATGGCCA	1748
QY	1933	CAGCCCTGAACAGAGAAATCATCTATGTTGTTAGAAATAGTAGACATTTTTTGGAAATG	1992
DB	1749	CAGCCCTGAACAGAGAAATCATCTATGTTGTTAGAAATAGTAGACATTTTTTGGAAATG	1808
QY	1993	TCCAGATGCGAGAGAGCGCTGTTTATGTTGCAATTTCCAGAGGCGGAAATGAGAGCGAAA	2052

Db 1809 TCGGAAGTCGACAGAGCGCTGTCTATTGGCAATTCAGAGCGGAAATGAAGAGCGAAA 1868
QY 2053 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGCTAG 2112
Db 1869 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGCTAG 1928
QY 2113 TCTACAACAGAGGATTCAGGCAATTAACCTCTGCCATGGGTGGAAACATGGGTTCATACA 2172
Db 1929 TCTACAACAGAGGATTCAGGCAATTAACCTCTGCCATGGGTGGAAACATGGGTTCATACA 1988
QY 2173 AACTCTTCTTAAGTAACTGGAAGTCAATTCAGACAGAGCATTTGGAAGAACTTCTTCA 2232
Db 1989 AACTCTTCTTAAGTAACTGGAAGTCAATTCAGACAGAGCATTTGGAAGAACTTCTTCA 2048
QY 2233 TAAAGATGATGAGGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2292
Db 2049 TAAAGATGATGAGGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2108
QY 2293 CCAGAGGTCGTGACAGAGACTTCATGAGCTCATCAACCAACCCCAATCTCAACAGCAT 2352
Db 2109 CCAGAGGTCGTGACAGAGACTTCATGAGCTCATCAACCAACCCCAATCTCAACAGCAT 2168
QY 2353 GGATGATGTTCTGTGAACAGTTTGGAAAGGACCGAAACAAAGCTCGGCAAGGCCAGG 2412
Db 2169 GGATGATGTTCTGTGAACAGTTTGGAAAGGACCGAAACAAAGCTCGGCAAGGCCAGG 2228
QY 2413 ACATACCCAGGAAACAGTAACAAATGGAAGCACTTTACAGAAATAAAGAAAGGTAGAAA 2472
Db 2229 ACATACCCAGGAAACAGTAACAAATGGAAGCACTTTACAGAAATAAAGAAAGGTAGAAA 2288
QY 2473 CAGGAGGACCGAGAAATTTGAGAGGACCCAGGAGTGTCTGAGTGCATTAACCTTAGA 2532
Db 2289 CAGGAGGACCGAGAAATTTGAGAGGACCCAGGAGTGTCTGAGTGCATTAACCTTAGA 2348
QY 2533 AACCTCAAAACAGTAGAAACTTGCCTAGACAAATACTGGAAACAAAGTGCATATACAT 2592
Db 2349 AACCTCAAAACAGTAGAAACTTGCCTAGACAAATACTGGAAACAAAGTGCATATACAT 2408
QY 2593 GAACCTTTTTCATGGCAATATGAGATGTTTACAATGGTGGGAAATTCAGCTGAGTTCCA 2652
Db 2409 GAACCTTTTTCATGGCAATATGAGATGTTTACAATGGTGGGAAATTCAGCTGAGTTCCA 2468
QY 2653 CCAATTAATAATTAATCCATGAGTAACTTCTCTAATAGGCTTTTTTTCC 2702
Db 2469 CCAATTAATAATTAATCCATGAGTAACTTCTCTAATAGGCTTTTTTTTC 2518

RESULT 9

US-08-488-935-1

; Sequence 1, Application US/08488935

; GENERAL INFORMATION:

; APPLICANT: Goodman, Corey S.

; APPLICANT: Kolodkin, Alex L.

; APPLICANT: Matthes, David

; APPLICANT: Bentley, David R.

; APPLICANT: O'Connor, Timothy

; TITLE OF INVENTION: Semaphorins

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: RICHARD ARON OSMAN, Ph.D.

; STREET: 268 Bush Street, Suite 3200

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/488,935

; FILING DATE:

; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/489,057
; FILING DATE: 09-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-2/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2601 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..2331
; US-08-488-935-1

Query Match

Best Local Similarity 92.6%; Score 2508.4; DB 9; Length 2601;

Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 193 CTGCAGCATGGCTGGTTAACTAGGATTTGCTCTCTTTCTGGGAGATTTACTTACAGC 252
Db 9 CTGCAGCATGGCTGGTTAACTAGGATTTGCTCTCTTTCTGGGAGATTTACTTACAGC 68
QY 253 AAGACAACTATCAGAAATGGGAAGAAACAATGTGCCAAGGCTGAAATATCTTACAAAGA 312
Db 69 AAGACAACTATCAGAAATGGGAAGAAACAATGTGCCAAGGCTGAAATATCTTACAAAGA 128
QY 313 AATGTTGGAATCCAAACAATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCA 372
Db 129 AATGTTGGAATCCAAACAATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCA 188
QY 373 TACCTTCCTTTTGGATGAGGAACGAGTAGCTGTATGTTGGAGCAAAAGGATCAATATT 432
Db 189 TACCTTCCTTTTGGATGAGGAACGAGTAGCTGTATGTTGGAGCAAAAGGATCAATATT 248
QY 433 TTCATTGACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTTGGCCAGTATCTTACAC 492
Db 249 TTCATTGACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTTGGCCAGTATCTTACAC 308
QY 493 CAGAAGAGATGAATGCAAGTGGGCTGGAAAGACATCTCTGAGGACAAATTTGCTTAATTTCA 552
Db 309 CAGAAGAGATGAATGCAAGTGGGCTGGAAAGACATCTCTGAGGACAAATTTGCTTAATTTCA 368
QY 553 CAAGTACTTAAGGCATATAATCAGACTCACTTGTACGCTGTGGAACGGGGCTTTTCA 612
Db 369 CAAGTACTTAAGGCATATAATCAGACTCACTTGTACGCTGTGGAACGGGGCTTTTCA 428
QY 613 TCCAAATTTGACCTACATTTGAAATTTGACATCATCTCTGAGGACAAATTTTAAAGCTGGA 672
Db 429 TCCAAATTTGACCTACATTTGAAATTTGACATCATCTCTGAGGACAAATTTTAAAGCTGGA 488
QY 673 GAATCACAATTTGAAAACCGGCGGTGGGAAGAGTCCATATGACCCCTAAGCTGTGACAGC 732
Db 489 GAATCACAATTTGAAAACCGGCGGTGGGAAGAGTCCATATGACCCCTAAGCTGTGACAGC 548
QY 733 ATCCCTTTAATAGATGGAGAAATTTACTCTGGAAGTGCAGCTGATTTTATGGGGCGAGA 792
Db 549 ATCCCTTTAATAGATGGAGAAATTTACTCTGGAAGTGCAGCTGATTTTATGGGGCGAGA 608
QY 793 CTTTGTCTATCTTCGAACTCTTTGGGCAACACCCCAATCAGGACAGAGCAGCATGATTC 852
Db 609 CTTTGTCTATCTTCGAACTCTTTGGGCAACACCCCAATCAGGACAGAGCAGCATGATTC 668
QY 853 CAGGTGGCTCAATGATCCAAAGTTTCAATAGTGGCCACCTCTATCTCAGAGAGTGACAATCC 912

Db 669 CAGGTGGCTCAATGATCCAAAGTTTCATTTAGTGGCCACCTCATCTCAGAGAGTGACAATCC 728
Qy 913 TGAAGATGCAAGATATACCTTTTCTTCGTTGAATGCAATAGATGGAGACACTCTGG 972
Db 729 TGAAGATGCAAGATATACCTTTTCTTCGTTGAATGCAATAGATGGAGACACTCTGG 788
Qy 973 AAAAGCTACTCAGCTAGAGATAGGTGAGATATGCAAGATGACTTTGGAGGGCACAGAAG 1032
Db 789 AAAAGCTACTCAGCTAGAGATAGGTGAGATATGCAAGATGACTTTGGAGGGCACAGAAG 848
Qy 1033 TCTGGTGAATAATGACCAACATTTCTCAAGCTCGTCTGATTTGCTCAGTGCCAGGTCC 1092
Db 849 TCTGGTGAATAATGACCAACATTTCTCAAGCTCGTCTGATTTGCTCAGTGCCAGGTCC 908
Qy 1093 AAATGSCATGACACTCACTTTTGTATGATGAATGCAAGATGACTTTGCAAGATGACTTTTAAAGA 1152
Db 909 AAATGSCATGACACTCACTTTTGTATGATGAATGCAAGATGACTTTTAAAGA 968
Qy 1153 TCCTAAATAATCCAGTTGTATATGAGTGTTCAGACTTCCAGTAACATTTTCAAGGGATC 1212
Db 969 TCCTAAATAATCCAGTTGTATATGAGTGTTCAGACTTCCAGTAACATTTTCAAGGGATC 1028
Qy 1213 AGCCGTGTGTATGATPAGCATGATGATGAGAGGGTGTTCCTTGGTCCATATGCCCA 1272
Db 1029 AGCCGTGTGTATGATPAGCATGATGATGAGAGGGTGTTCCTTGGTCCATATGCCCA 1088
Qy 1273 CAGGGATGACCAACTATCAATGGTGTCTTATCAAGAGAGTCCCTATCCAGGGC 1332
Db 1089 CAGGGATGACCAACTATCAATGGTGTCTTATCAAGAGAGTCCCTATCCAGGGC 1148
Qy 1333 AGGAACCTGTCCAGCAAAACATTTGGTGTGTGACTCTCAAGAGGACTTTCCTGATGA 1392
Db 1149 AGGAACCTGTCCAGCAAAACATTTGGTGTGTGACTCTCAAGAGGACTTTCCTGATGA 1208
Qy 1393 TGTATPAACTTTTGAAGAGTCAATCCAGCCATGTACAAATCAGATGTTTCTTATGAACAA 1452
Db 1209 TGTATPAACTTTTGAAGAGTCAATCCAGCCATGTACAAATCAGATGTTTCTTATGAACAA 1268
Qy 1453 TCGCCCAATAGTATCAAAACCGATGTAATTTATCAATTTACAAATTTGTCGTAGACCG 1512
Db 1269 TCGCCCAATAGTATCAAAACCGATGTAATTTATCAATTTTACAAATTTGTCGTAGACCG 1328
Qy 1513 AGTGGATGCAAGATGAGACAGTATGATGTTATGTTTATCGAAACAGATGTTGGAGCGT 1572
Db 1329 AGTGGATGCAAGATGAGACAGTATGATGTTATGTTTATCGAAACAGATGTTGGAGCGT 1388
Qy 1573 TCTTAAAGTATGTTCAATTTCTTAAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGA 1632
Db 1389 TCTTAAAGTATGTTCAATTTCTTAAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGA 1448
Qy 1633 AGNAATGACAGTTTTCGGGAACCGACTGTATTTTCCAGCAATGGAGCTTTCACCTAAGCA 1692
Db 1449 AGNAATGACAGTTTTCGGGAACCGACTGTATTTCCAGCAATGGAGCTTTCACCTAAGCA 1508
Qy 1693 GCAACAACTATATATTTGTTTCAACGCTGCGGGTTCAGCTCCCTTTACACGGTGTGA 1752
Db 1509 GCAACAACTATATATTTGTTTCAACGCTGCGGGTTCAGCTCCCTTTACACGGTGTGA 1568
Qy 1753 TATTTACGGGAAGCGTGTGCTGAGTGTTCCTTCGCGGAGACCTTACTGTCTTGGGA 1812
Db 1569 TATTTACGGGAAGCGTGTGCTGAGTGTTCCTTCGCGGAGACCTTACTGTCTTGGGA 1628
Qy 1813 TGGTCTGATGTTCTGCTATTTTCCACTGCAAGAGACCGCACAGACGACAGATAT 1872
Db 1629 TGGTCTGATGTTCTGCTATTTTCCACTGCAAGAGACCGCACAGACGACAGATAT 1688
Qy 1873 AGNAATGAGAGCCCACTGACTCACTGTTTCAAGCTTACACCATGATAATCACCATGGCCA 1932
Db 1689 AGNAATGAGAGCCCACTGACTCACTGTTTCAAGCTTACACCATGATAATCACCATGGCCA 1748
Qy 1933 CAGCCCTGAAGAGAGAAATCATCTATGGTGTAGAGATAGTACACATTTTTCGAATGCGAG 1992

Db 1749 CAGCCCTGAAGAGAGAAATCATCTATGTGTAGAGAATAGTACACATTTTTCGAATGCGAG 1808
Qy 1933 TCCGAAGTCGAGAGAGCGCTGTCTATTTGGCAATTTCCAGAGGCGAAATGAAGAGCGAAA 2052
Db 1809 TCCGAAGTCGAGAGAGCGCTGTCTATTTGGCAATTTCCAGAGGCGAAATGAAGAGCGAAA 1868
Qy 2053 AGAAGAGATCAGAGTGGATGATCATATCATCATCAGGACAGATCAAGGCCCTTCTGCTACGTTAG 2112
Db 1869 AGAAGAGATCAGAGTGGATGATCATATCATCATCAGGACAGATCAAGGCCCTTCTGCTACGTTAG 1928
Qy 2113 TCTCAACAGAGAGATTCAGGCAATTCCTCTGCCATCGGTGGAGAACATGGGTTTCATACA 2172
Db 1929 TCTCAACAGAGAGATTCAGGCAATTCCTCTGCCATCGGTGGAGAACATGGGTTTCATACA 1988
Qy 2173 AACTCTTCTTAAGGTAACCTCGGAAGTCAATGACACAGAGCAATTTGGAGAACTTCTTCA 2232
Db 1989 AACTCTTCTTAAGGTAACCTCGGAAGTCAATGACACAGAGCAATTTGGAGAACTTCTTCA 2048
Qy 2233 TAAAGATGATGATGGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2292
Db 2049 TAAAGATGATGATGGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2108
Qy 2293 CCAGAGGTCTGGTACAGAGACTTTCATGAGAGTCTATCAACCCCAATCTCAACAGAT 2352
Db 2109 CCAGAGGTCTGGTACAGAGACTTTCATGAGAGTCTATCAACCCCAATCTCAACAGAT 2168
Qy 2353 GATGAGTCTGTGAACAAAGTTTGGAAAGGAGCCGAAACAAACGTCGCGAAAGGCCAGG 2412
Db 2169 GATGAGTCTGTGAACAAAGTTTGGAAAGGAGCCGAAACAAACGTCGCGAAAGGCCAGG 2228
Qy 2413 ACATACCCCGAGGAAACAGTAACAAATGGAAGCACTTACAGAAAAATAAGAAAGGTAGAAA 2472
Db 2229 ACATACCCCGAGGAAACAGTAACAAATGGAAGCACTTACAGAAAAATAAGAAAGGTAGAAA 2288
Qy 2473 CAGAGGACCCACGAATTTGAGAGGACCCAGAGAGTGTCTGAGCTGATTAACCTCTAGA 2532
Db 2289 CAGAGGACCCACGAATTTGAGAGGACCCAGAGAGTGTCTGAGCTGATTAACCTCTAGA 2348
Qy 2533 AACCTCAACAAAGTAGAAACTTTCCTACACAAATTAACCTGAGAAACAAATGCAATATACAT 2592
Db 2349 AACCTCAACAAAGTAGAAACTTTCCTACACAAATTAACCTGAGAAACAAATGCAATATACAT 2408
Qy 2593 GAACTTTTTCATGGCATTATGTGGATGTTTACAAATGGTGGGAAATTCAGCTGAGTTCCA 2652
Db 2409 GAACTTTTTCATGGCATTATGTGGATGTTTACAAATGGTGGGAAATTCAGCTGAGTTCCA 2468
Qy 2653 CCAATTAATAATTAATCCATGAGTAACCTTTCCTTAATAGGCTTTTTTTC 2702
Db 2469 CCAATTAATAATTAATCCATGAGTAACCTTTCCTTAATAGGCTTTTTTTC 2518

RESULT 10
US-08-489-057-1
; Sequence 1, Application US/08489057
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David R.
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: Semaphorins
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RICHARD ARON OSMAN, Ph.D.
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS

Db 1749 CAGCCCTGAAGAGAGAAATCATCTATGCTGTAGAGAAATAGTACACATTTTGGAAATGCAG 1808
Qy 1993 TCCGAAGTCCGAGAGAGCCCTGGTCTATTGGCAATTCAGAGGCGGAATGAAGAGCGAA 2052
Db 1809 TCCGAAGTCCGAGAGAGCCCTGGTCTATTGGCAATTCAGAGGCGGAATGAAGAGCGAA 1868
Qy 2053 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGCTAG 2112
Db 1869 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGCTAG 1928
Qy 2113 TCTACAACAGAGAGATTCAGGCAATTCCTCTGCCATGCGGTGGAAATGCGGTTCATACA 2172
Db 1929 TCTACAACAGAGAGATTCAGGCAATTCCTCTGCCATGCGGTGGAAATGCGGTTCATACA 1988
Qy 2173 AACTCTTCTTAAGTTAACTGAGTCAATGACACAGAGCAATTTGGAGAACTTCTTCA 2232
Db 1989 AACTCTTCTTAAGTTAACTGAGTCAATGACACAGAGCAATTTGGAGAACTTCTTCA 2048
Qy 2233 TAAAGATGATGAGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2292
Db 2049 TAAAGATGATGAGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2108
Qy 2293 CCAGAGAGTCTGGTACAGAGACTTCATGAGTCTCATCAACCCCAATCTCAACACGAT 2352
Db 2109 CCAGAGAGTCTGGTACAGAGACTTCATGAGTCTCATCAACCCCAATCTCAACACGAT 2168
Qy 2353 GGATGAGTTCTGTGAACAGTTTGGAAAGGACCGAAACCAACGTCGCAAGGCCAGG 2412
Db 2169 GGATGAGTTCTGTGAACAGTTTGGAAAGGACCGAAACCAACGTCGCAAGGCCAGG 2228
Qy 2413 ACATACCCAGGACAGTAAACAAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAA 2472
Db 2229 ACATACCCAGGACAGTAAACAAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAA 2288
Qy 2473 CAGGAGGCCACGAAATTTGAGAGGCCACCCAGGAGTGTCTGAGCTGCAATTCCTCTAGA 2532
Db 2289 CAGGAGGCCACGAAATTTGAGAGGCCACCCAGGAGTGTCTGAGCTGCAATTCCTCTAGA 2348
Qy 2533 AACCTCAACAGTAGAAACTTGCCTAGACAACTGCAAAACCAATGCAATATACAT 2592
Db 2349 AACCTCAACAGTAGAAACTTGCCTAGACAACTGCAAAACCAATGCAATATACAT 2408
Qy 2593 GAACTTTTTCATGGCATTATGTGGATGTTTCAATGGTGGGAAATTCAGCTGAGTTCCA 2652
Db 2409 GAACTTTTTCATGGCATTATGTGGATGTTTCAATGGTGGGAAATTCAGCTGAGTTCCA 2468
Qy 2653 CCAATTAATAATTAATCCATGAGTAACTTCTTAATAGGCTTTTTCCTCC 2702
Db 2469 CCAATTAATAATTAATCCATGAGTAACTTCTTAATAGGCTTTTTCCTCC 2518

RESULT 11
US-10-067-632-53
; Sequence 53, Application US/10067632
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; Kolodkin, Alex L.
; Matthes, David
; Bentley, David R.
; O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/067,632
FILING DATE: 04-Feb-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/060,610
FILING DATE: <Unknown>
APPLICATION NUMBER: 08/835,268
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 2601 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 16...2331
SEQUENCE DESCRIPTION: SEQ ID NO: 53:
US-10-067-632-53
Query Match 92.6%; Score 2508.4; DB 45; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 193 CTGCAGCATGGCTGGTTAACTAGGATTTGTCTCTTTTCTGGGAGTATTACTTACAGC 252
Db 9 CTGCAGCATGGCTGGTTAACTAGGATTTGTCTCTTTTCTGGGAGTATTACTTACAGC 68
Qy 253 AAGAGCAACTATCAGAAATGGAGAGCAATGTGCCAAGGCTGAATTTATCTTACAAAGA 312
Db 69 AAGAGCAACTATCAGAAATGGAGAGCAATGTGCCAAGGCTGAATTTATCTTACAAAGA 128
Qy 313 AATGTTGGAATCCAAATGTGATCCTTTCAATGGCTTGGCCAAAGCTCCAGTTTATCA 372
Db 129 AATGTTGGAATCCAAATGTGATCCTTTCAATGGCTTGGCCAAAGCTCCAGTTTATCA 188
Qy 373 TACCTTCCTTTTGGATGAGGAAACGAGTAGGCTGTATGTTGGAGCAAGGATCAATAT 432
Db 189 TACCTTCCTTTTGGATGAGGAAACGAGTAGGCTGTATGTTGGAGCAAGGATCAATAT 248
Qy 433 TTCAATTCAGCTGGTTAATATCAAGGATTTTCAAGAGATTTGTGGCCAGTATCTTACAC 492
Db 249 TTCAATTCAGCTGGTTAATATCAAGGATTTTCAAGAGATTTGTGGCCAGTATCTTACAC 308
Qy 493 CAGAGAGATCAATGCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTGCTAAATTTTCA 552
Db 309 CAGAGAGATCAATGCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTGCTAAATTTTCA 368
Qy 553 CAAGGTACTTAAGGCATATAATCAGACTCCTTGTACCCCTGTGGAAACGCGGCTTTTCA 612
Db 369 CAAGGTACTTAAGGCATATAATCAGACTCCTTGTACCCCTGTGGAAACGCGGCTTTTCA 428
Qy 613 TCCAAATTTGACCTACATTTGAAATTTGGACATCTCTGAGGACATATTTTAAAGCTGGA 672
Db 429 TCCAAATTTGACCTACATTTGAAATTTGGACATCTCTGAGGACATATTTTAAAGCTGGA 488
Qy 673 GAACTCACATTTTGAAGAGCGGCGGAGAGAGTCCATATGACCCCTAGCTCTGACAGC 732
Db 489 GAACTCACATTTTGAAGAGCGGCGGAGAGAGTCCATATGACCCCTAGCTCTGACAGC 548
Qy 733 ATCCCTTTTAATAGATGGAGAAATATATCTCTGGAACCTGAGCTGATTTTATGGGGGAGA 792

Db 549 ATCCCTTTTAAATAGATGGAGAAATATATCTCTGGAATCGAGCTGATTTTATGGGCGAGA 608
Qy 793 CTTTGTATCTTCCGAATCTTTGGGCAACCAACCAATCAGGACAGACGATGATTC 852
Db 609 CTTTGTATCTTCCGAATCTTTGGGCAACCAACCAATCAGGACAGACGATGATTC 668
Qy 853 CAGGTGGCTCAATGATCCAAAGTTCAATAGTCCCACTCTCATCTCAGAGAGTGACAATCC 912
Db 669 CAGGTGGCTCAATGATCCAAAGTTCAATAGTCCCACTCTCATCTCAGAGAGTGACAATCC 728
Qy 913 TGAAGATGACAAAGTATATCTTTTCTCGTGAATGCAATAGATGGAGACACTCTGG 972
Db 729 TGAAGATGACAAAGTATATCTTTTCTCGTGAATGCAATAGATGGAGACACTCTGG 788
Qy 973 AAAAGCTACTCAGCTAGATAGGTGAGATATGCAAGATGACTTTGGAGGGCACAGAAG 1032
Db 789 AAAAGCTACTCAGCTAGATAGGTGAGATATGCAAGATGACTTTGGAGGGCACAGAAG 848
Qy 1033 TCTGTGTAATAATGACAACTCTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCC 1092
Db 849 TCTGTGTAATAATGACAACTCTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCC 908
Qy 1093 AAATGGCATGACACTATCTTTGATGAACTGAGGATGATTTCTTAATGAATTTAAAGA 1152
Db 909 AAATGGCATGACACTATCTTTGATGAACTGAGGATGATTTCTTAATGAATTTAAAGA 968
Qy 1153 TCTTAAATAATCCAGTGTATATGAGAGTGTATGAGACTTCCAGTAACTTTCAAGGGATC 1212
Db 969 TCTTAAATAATCCAGTGTATATGAGAGTGTATGAGACTTCCAGTAACTTTCAAGGGATC 1028
Qy 1213 AGCCGTGTGTATGATAGATGAGTGTATGAGAGGGTGTCTTGTGTCATATGCCCA 1272
Db 1029 AGCCGTGTGTATGATAGATGAGTGTATGAGAGGGTGTCTTGTGTCATATGCCCA 1088
Qy 1273 CAGGATGACCAACTATCAATGGTGGCTTATCAAGGAGATGCCCTATCCAGGCC 1332
Db 1089 CAGGATGACCAACTATCAATGGTGGCTTATCAAGGAGATGCCCTATCCAGGCC 1148
Qy 1333 AGCAACTTGTCCAGCAAAATTTGGTGTGTGACTCTCAAAAGGACCTTCTGATGA 1392
Db 1149 AGCAACTTGTCCAGCAAAATTTGGTGTGTGACTCTCAAAAGGACCTTCTGATGA 1208
Qy 1393 TCTTAAAGTAGTTCAAATCTTCAAGGACCTTGTGATGATTTAGAGAGGTTCCTGATGA 1452
Db 1209 TGTATTAACCTTTGCAAGAGTATCCAGCCATGATCAATCCAGTGTTCCTATGAACA 1268
Qy 1453 TCGCCCAATAGTATCAAAACCGATGTAATTTATCAATTTACAAATTTGCTGATGACCG 1512
Db 1269 TCGCCCAATAGTATCAAAACCGATGTAATTTATCAATTTACAAATTTGCTGATGACCG 1328
Qy 1513 AGTGGATGACAAAGATGACAGTATGATGATTTATCGGAAACAGATGTTGGGACCGT 1572
Db 1329 AGTGGATGACAAAGATGACAGTATGATGATTTATCGGAAACAGATGTTGGGACCGT 1388
Qy 1573 TCTTAAAGTAGTTCAAATCTTCAAGGACCTTGTGATGATTTAGAGAGGTTCCTGATGA 1632
Db 1389 TCTTAAAGTAGTTCAAATCTTCAAGGACCTTGTGATGATTTAGAGAGGTTCCTGATGA 1448
Qy 1633 AGAAATGACAGTTCCTGGGAAACCGATGATGATTTATCGGAAACAGATGTTCCACTAAGA 1692
Db 1449 AGAAATGACAGTTCCTGGGAAACCGATGATGATTTATCGGAAACAGATGTTCCACTAAGA 1508
Qy 1693 GCAACAACTATATTTGGTTCAACGGCTGGGGTGCAGCTCCCTTTACACCGGTGGA 1752
Db 1509 GCAACAACTATATTTGGTTCAACGGCTGGGGTGCAGCTCCCTTTACACCGGTGGA 1568
Qy 1753 TATTTACGGGAAAGCGTGTGCTGAGTGTTCCTCGCCCGAGACCTTACTGTGCTGGGA 1812
Db 1569 TATTTACGGGAAAGCGTGTGCTGAGTGTTCCTCGCCCGAGACCTTACTGTGCTGGGA 1628
Qy 1813 TGGTTCTGATGTTCTCGCTATTTTCCACTGCAAGAGACCGCAACAGCGCAAGATAT 1872

Db 1629 TGGTTCTGCATGTTCTCGCTATTTTCCCACTGCAAAAGAGACGCAACAGACGACAGATAT 1688
Qy 1873 AAGAAATGGAGACCCACTGACTCACTGTTTACAGCTTTACACCATGATAATCACCATTGGCCA 1932
Db 1689 AAGAAATGGAGACCCACTGACTCACTGTTTACAGCTTTACACCATGATAATCACCATTGGCCA 1748
Qy 1933 CAGCCCTCAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGCACAATTTTGGAAATGCGAG 1992
Db 1749 CAGCCCTCAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGCACAATTTTGGAAATGCGAG 1808
Qy 1993 TCCGAAATGCGAGAGAGCGCTGCTATTTGGCAATTTCCAGAGCGGCAAAATGAGAGCGGAAA 2052
Db 1809 TCCGAAATGCGAGAGAGCGCTGCTATTTGGCAATTTCCAGAGCGGCAAAATGAGAGCGGAAA 1868
Qy 2053 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAG 2112
Db 1869 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAG 1928
Qy 2113 TCTACAAAGAGAGTTCAGGCAATTAATCTCTGCTGATGAGGAAATGAGGTTTCAACA 2172
Db 1929 TCTACAAAGAGAGTTCAGGCAATTAATCTCTGCTGATGAGGAAATGAGGTTTCAACA 1988
Qy 2173 AACTCTTCTTAAGGTAAACCTTGGAAATTAATGACACAGAGCATTTGGAAAGACTTCTTCA 2232
Db 1989 AACTCTTCTTAAGGTAAACCTTGGAAATTAATGACACAGAGCATTTGGAAAGACTTCTTCA 2048
Qy 2233 TAAAGATGATGATGAGATGCTCTTAAGCAAAAGAAATGTCCAATAGCATCACACCTAG 2292
Db 2049 TAAAGATGATGATGAGATGCTCTTAAGCAAAAGAAATGTCCAATAGCATCACACCTAG 2108
Qy 2293 CCAGAGTCTGTTACAGAGATCTTCAAGCTCAATCAACCAACCCCAATCTCAACACGAT 2352
Db 2109 CCAGAGTCTGTTACAGAGATCTTCAAGCTCAATCAACCAACCCCAATCTCAACACGAT 2168
Qy 2353 GGATGAGTCTCTGTAACAAAGTTTGGAAAGGACCGGAAACCAACCTCGCAAGGCCAGG 2412
Db 2169 GGATGAGTCTCTGTAACAAAGTTTGGAAAGGACCGGAAACCAACCTCGCAAGGCCAGG 2228
Qy 2413 ACATACCCCGAGGACAGTAAATGGAAGCACTTACAGAAATAAGAAAGGTAGAAA 2472
Db 2229 ACATACCCCGAGGACAGTAAATGGAAGCACTTACAGAAATAAGAAAGGTAGAAA 2288
Qy 2473 CAGGAGGCCCAACGAATTTGAGAGGCCACCCAGGAGTCTCTGAGCTGCAATACCTCTAGA 2532
Db 2289 CAGGAGGCCCAACGAATTTGAGAGGCCACCCAGGAGTCTCTGAGCTGCAATACCTCTAGA 2348
Qy 2533 AACCTCAAAAGTAGAAACTTGCCTAGACAATTAAGTGAAGGAAACAAATGCATATACAT 2592
Db 2349 AACCTCAAAAGTAGAAACTTGCCTAGACAATTAAGTGAAGGAAACAAATGCATATACAT 2408
Qy 2593 GAACTTTTTCATGGCATTTATGAGTCTTTTACAAATGTTGGGAAATTCAGCTGAGTTCCA 2652
Db 2409 GAACTTTTTCATGGCATTTATGAGTCTTTTACAAATGTTGGGAAATTCAGCTGAGTTCCA 2468
Qy 2653 CCAATTAATAATTAATCCATGAGTAACTTCTTCAATAGGCTTTTTTTC 2702
Db 2469 CCAATTAATAATTAATCCATGAGTAACTTCTTCAATAGGCTTTTTTTC 2518

RESULT 12

US-10-170-235-19890

; Sequence 19890, Application US/10170235

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig

; TITLE OF INVENTION: KITS, SUCH AS NUCLEIC ACID ARRAYS, COMPRISING A MAJORITY OF HUMEN

; FILE REFERENCE: CL001380

; CURRENT APPLICATION NUMBER: US/10/170,235

; CURRENT FILING DATE: 2003-03-17

; NUMBER OF SEQ ID NOS: 42514

; SEQ ID NO 19890

; LENGTH: 2857

; TYPE: DNA

ORGANISM: HUMAN
US-10-170-235-19890

Query Match 92.4%; Score 2502.4; DB 47; Length 2857;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2516; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 193 CTGCAGCATGGGCTGGTAACTAGGATGTCGTCTTTCTGGGAGTAGTACTTACAGC 252
DB 1 CTGCAGCATGGGCTGGTAACTAGGATGTCGTCTTTCTGGGAGTAGTACTTACAGC 60

QY 253 AAGAGCAAACTATCAGAAATGGGAAGAACAAATGCGCAAGGCTCAAAATATCTTACAAAGA 312
DB 61 AAGAGCAAACTATCAGAAATGGGAAGAACAAATGCGCAAGGCTCAAAATATCTTACAAAGA 120

QY 313 AATGTTGGAATCCAAATATGATACATCTTTCAATGGCTTGGCCCAACAGCTCCAGTTATCA 372
DB 121 AATGTTGGAATCCAAATATGATACATCTTTCAATGGCTTGGCCCAACAGCTCCAGTTATCA 180

QY 373 TACCTTCCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATATT 432
DB 181 TACCTTCCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATATT 240

QY 433 TTCAATTCGACCTGGTTAAATPATCAAGGATTTTCAAAAGATTTGTGTGGCCAGTATCTTACAC 492
DB 241 TTCAATTCGACCTGGTTAAATPATCAAGGATTTTCAAAAGATTTGTGTGGCCAGTATCTTACAC 300

QY 493 CAGAGAGATGAATGCAAGTGGCTGGGCGGAAAGACATCTCTGAAAGATGTGCTAAATTTCA 552
DB 301 CAGAGAGATGAATGCAAGTGGCTGGGCGGAAAGACATCTCTGAAAGATGTGCTAAATTTCA 360

QY 553 CAAGGTACTTAAAGGCATATAATCAGACTCACTTGTACGCTCTGGAACGGGGCTTTTCA 612
DB 361 CAAGGTACTTAAAGGCATATAATCAGACTCACTTGTACGCTCTGGAACGGGGCTTTTCA 420

QY 613 TCCAAATTCGACCTAATGAAATGGACATCATCTGAGGACAATATTTTAAAGCTGGA 672
DB 421 TCCAAATTCGACCTAATGAAATGGACATCATCTGAGGACAATATTTTAAAGCTGGA 480

QY 673 GAACTCACATTTTGAACCGGCGTGGGAAGAGTCCATATGACCCTTAAGCTCTCAGACG 732
DB 481 GAACTCACATTTTGAACCGGCGTGGGAAGAGTCCATATGACCCTTAAGCTCTCAGACG 540

QY 733 ATCCCTTTTAAATAGATGGAGAAATTAATCTCTGAACTGCACTGATTTTATGGGCGGAGA 792
DB 541 ATCCCTTTTAAATAGATGGAGAAATTAATCTCTGAACTGCACTGATTTTATGGGCGGAGA 600

QY 793 CTTTGTATCTTCCGAACTCTTTGGCACCACCACCAATCAGGACAGACGATGATTC 852
DB 601 CTTTGTATCTTCCGAACTCTTTGGCACCACCACCAATCAGGACAGACGATGATTC 660

QY 853 CAGGTGGCTCAATGATCCAAAGTTCAATAGTCCACCTCTCTCAGAGAGTGAATCC 912
DB 661 CAGGTGGCTCAATGATCCAAAGTTCAATAGTCCACCTCTCTCAGAGAGTGAATCC 720

QY 913 TGAAGATGACAAAGTATACCTTTCTCGTGAATGCAATAGATGAGAGACATCTGG 972
DB 721 TGAAGATGACAAAGTATACCTTTCTCGTGAATGCAATAGATGAGAGACATCTGG 780

QY 973 AAAAGCTACTCAGCTAGAAATAGGTGAGATATGCAAGATGATCTTTGGAGGSCACAGAAG 1032
DB 781 AAAAGCTACTCAGCTAGAAATAGGTGAGATATGCAAGATGATCTTTGGAGGSCACAGAAG 840

QY 1033 TCTGGTGAATAATGGACAACTTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCC 1092
DB 841 TCTGGTGAATAATGGACAACTTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCC 900

QY 1093 AAATGGCATTGACCTCACTTTTGAAGTGAAGTATGATCTTAAATGAACCTTTAAAGA 1152
DB 901 AAATGGCATTGACCTCACTTTTGAAGTGAAGTATGATCTTAAATGAACCTTTAAAGA 960

QY 1153 TCCTAAAAATCCAGTTGTATAGGATGTTTACGACTTCCAGTAACTTTTCAAGGGATC 1212
DB TCCTAAAAATCCAGTTGTATAGGATGTTTACGACTTCCAGTAACTTTTCAAGGGATC 2100

DB 961 TCCTAAAAATCCAGTTGTATAGGATGTTTACGACTTCCAGTAACTTTTCAAGGGATC 1020

QY 1213 AGCGGTGTATGTATAGCATGATGATGAGAGGGGTGTTCTTGGTCCATATGCCCCA 1272

DB 1021 AGCGGTGTATGTATAGCATGATGATGAGAGGGGTGTTCTTGGTCCATATGCCCCA 1080

QY 1273 CAGGATCGGACCCAACTACTCAATGGGTGCTTATCAAGGAAGAGTCCCTTATCCAGGCC 1332

DB 1081 CAGGATCGGACCCAACTACTCAATGGGTGCTTATCAAGGAAGAGTCCCTTATCCAGGCC 1140

QY 1333 AGGAATCTGTCCCAGCAAAACATTTGGTGGTTTGACTCTCAAAAGGACCTTCTGATGA 1392

DB 1141 AGGAATCTGTCCCAGCAAAACATTTGGTGGTTTGACTCTCAAAAGGACCTTCTGATGA 1200

QY 1393 TGTATTAACCTTTGCAAGAGTCAATCCAGCCATGTACAAATCCAGTGTTCCTATGAACA 1452

DB 1201 TGTATTAACCTTTGCAAGAGTCAATCCAGCCATGTACAAATCCAGTGTTCCTATGAACA 1260

QY 1453 TCGCCCAATAGTGTCAAAACGGATGTAAATTTATCAATTTTACAAAATTTGCTAGACCG 1512

DB 1261 TCGCCCAATAGTGTCAAAACGGATGTAAATTTATCAATTTTACAAAATTTGCTAGACCG 1320

QY 1513 AGTGATGCAAGATGACATGATGATGTTATGTTATCGGAAACAGATGTTGGACCGT 1572

DB 1321 AGTGATGCAAGATGACATGATGATGTTATGTTATCGGAAACAGATGTTGGACCGT 1380

QY 1573 TCTTAAAGTAGTTTCAATTTCCCTAAGGACATGTTGATGATTTAGGAAGAGTTCCTCTGA 1632

DB 1381 TCTTAAAGTAGTTTCAATTTCCCTAAGGACATGTTGATGATTTAGGAAGAGTTCCTCTGA 1440

QY 1633 AGAATGACATGTTTTCGGGAACCGACTGCTATTTTACGAAATGGAAGTTCCTCACTA---A 1689

DB 1441 AGAATGACATGTTTTCGGGAACCGACTGCTATTTTACGAAATGGAAGTTCCTCACTAAGA 1500

QY 1690 CGAGCAACAACTATATATTTGTTCAAAGTGGGGTTCGCCAGCTTCCCTTACACGGTG 1749

DB 1501 CGAGCAACAACTATATATTTGTTCAAAGTGGGGTTCGCCAGCTTCCCTTACACGGTG 1560

QY 1750 TGATATTTACGGGAAGCGTGTGAGTGTGCTCGCCGAGACCTTACTGTGCTTG 1809

DB 1561 TGATATTTACGGGAAGCGTGTGAGTGTGCTCGCCGAGACCTTACTGTGCTTG 1620

QY 1810 GGATGTTTCTCATGTTCTCGCTATTTTCCCACTGCAAGAGAGCGCAACAAGCAAGA 1869

DB 1621 GGATGTTTCTCATGTTCTCGCTATTTTCCCACTGCAAGAGAGCGCAACAAGCAAGA 1680

QY 1870 TATAAGAAATGGAGACCCACTGACTCTGTTTCACTTACCATGATTAATCACCATGG 1929

DB 1681 TATAAGAAATGGAGACCCACTGACTCTGTTTCACTTACCATGATTAATCACCATGG 1740

QY 1930 CCAGAGCCCTCAAGAGAGAAATCATCTATGGTGTAGAGAAATAGTAGCAATTTTGGAAATG 1989

DB 1741 CCAGAGCCCTCAAGAGAGAAATCATCTATGGTGTAGAGAAATAGTAGCAATTTTGGAAATG 1800

QY 1990 CAGTCCGAAGTCGAGAGAGCGCTGTTCTATTTGGCAATTTCCAGAGCGGCAATGAAGCGG 2049

DB 1801 CAGTCCGAAGTCGAGAGAGCGCTGTTCTATTTGGCAATTTCCAGAGCGGCAATGAAGCGG 1860

QY 2050 AAAAGAGAGATCAGAGTGGATGATCATCATCAGGACAGATCAAGGCCCTTCTGCTACG 2109

DB 1861 AAAAGAGAGATCAGAGTGGATGATCATCATCAGGACAGATCAAGGCCCTTCTGCTACG 1920

QY 2110 TAGTCTCAACAGAGAGGATTCAGGCAATTAATCTCTGCCATCGGTGGAAACATGGGTTCTAT 2169

DB 1921 TAGTCTCAACAGAGAGGATTCAGGCAATTAATCTCTGCCATCGGTGGAAACATGGGTTCTAT 1980

QY 2170 ACAAACCTCTTCTTAAAGGTAAACCTCGGAAGTCAATTTGACAAGAGTTCCTTGGAGAACTTCT 2229

DB 1981 ACAAACCTCTTCTTAAAGGTAAACCTCGGAAGTCAATTTGACAAGAGTTCCTTGGAGAACTTCT 2040

QY 2230 TCATAAAGATGATGATGGAGATGGCTCTAAGACCAAGAAATGTCCAAATAGCATGACACC 2289

DB 2041 TCATAAAGATGATGATGGAGATGGCTCTAAGACCAAGAAATGTCCAAATAGCATGACACC 2100

QY 2290 TAGCCAGAGAGGCTGTGTACAGAGACTTTCATGAGCTCATCAACCCCAATCTCAACAC 2349
DB |||||
DB 2101 TAGCCAGAGAGGCTGTGTACAGAGACTTTCATGAGCTCATCAACCCCAATCTCAACAC 2160
QY 2350 GATGATGAGTGTCTGTGAAACAGTGTGGAAGAGGACCAAAACAAACAGCTGCGCAAGGCC 2409
DB |||||
DB 2161 GATGATGAGTGTCTGTGAAACAGTGTGGAAGAGGACCAAAACAAACAGCTGCGCAAGGCC 2220
QY 2410 AGGACATACCCAGGGAACAGTAACAAATGGAAGCACTTACAGAAATAAGAAAGGTAG 2469
DB |||||
DB 2221 AGGACATACCCAGGGAACAGTAACAAATGGAAGCACTTACAGAAATAAGAAAGGTAG 2280
QY 2470 AAACAGGAGACCCAGCAATTTGAGAGGACCCAGGAGTGTCTGAGCTGCAATACCTCT 2529
DB |||||
DB 2281 AAACAGGAGACCCAGCAATTTGAGAGGACCCAGGAGTGTCTGAGCTGCAATACCTCT 2340
QY 2530 AGAAACCTCAAAACAGTACAAATTTGCCCTAGACAATACTGGAAACAAATGCAATATA 2589
DB |||||
DB 2341 AGAAACCTCAAAACAGTACAAATTTGCCCTAGACAATACTGGAAACAAATGCAATATA 2400
QY 2590 CATGAACCTTTTTCATGGCAATATATGGAATGTTTCAATGGTGGGAAATTCAGCTGAGTT 2649
DB |||||
DB 2401 CATGAACCTTTTTCATGGCAATATATGGAATGTTTCAATGGTGGGAAATTCAGCTGAGTT 2460
QY 2650 CCACCAATTAATAATTAATTCATGAGTAACCTTTCTTAATAGCTTTTTCCTAATACC 2709
DB |||||
DB 2461 CCACCAATTAATAATTAATTCATGAGTAACCTTTCTTAATAGCTTTTTCCTAATACC 2520

RESULT 13
US-10-144-771-9178
; Sequence 9178, Application US/10144771
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: HUMAN GENOME DISCOVERY SYSTEM AND USES THEREOF
; FILE REFERENCE: CL001321
; CURRENT APPLICATION NUMBER: US/10/144, 771
; CURRENT FILING DATE: 2002-05-15
; NUMBER OF SEQ ID NOS: 47235
; SEQ ID NO 9178
; LENGTH: 2913
; TYPE: DNA
; ORGANISM: HUMAN
US-10-144-771-9178

Query Match 75.7%; Score 2050; DB 46; Length 2913;
Best Local Similarity 88.1%; Pred. No. 0;
Matches 2281; Conservative 0; Mismatches 295; Indels 14; Gaps 4;

QY 118 TTTCCAGATGTTTGAACCTTCTCTGGCGGCAACAATACAGGAAGGAAGACTTAAAGCAGCAA 177
DB |||||
DB 29 TTTCCAGATGTTTGAACCTTCTCTGGCGGCAACAATACAGGAAGGAAGCTTCCGCGAGCTC 88
QY 178 AGGACCTTACAGCGCTGTGAGCATGGGCTGGTTAACTAGGATGTCTGTCTTTCTGGGG 237
DB |||||
DB 89 AGGACCTTACAGCGCTGTGAGCATGGGCTGGTTAACTAGGATGTCTGTCTTTCTGGGG 148
QY 238 AGTATTACTTACAGCAAGAGCAACTATCAGATGGGAAGCAATGTCCCAAGGCTGAA 297
DB |||||
DB 149 TGTATTACTTACAGCAGAGCAACTATGCAAAACGGAAGCAATGTGCCAAGACTGAA 208
QY 298 ATTATCTCAAAAGAAATGTTGGAATCCAAATGTGATCACTTTCAATGGCTGGCCAA 357
DB |||||
DB 209 ATTATCGTACAAAGAAATGTTGGAATCCAAATGTGATCACTTTTATGGCTGGCCAA 268
QY 358 CAGCTCCAGTTATCATACCTTCTCTTTTGGATGAGAAACGAGTAGGCTGTATGTTGGAGC 417
DB |||||
DB 269 CAGCTCCAGTTACCAACCTTCTCTCTTTGGATGAGAAACGAGTAGACTATATGTTGGAGC 328
QY 418 AAAGGATCACATTTTTCATTCGACCTGGTTAATATCAAGGATTTTCAAAGATGTTGTG 477
DB |||||
DB 329 AAAAGATCATATTTTTCATTCCTGTTGAACATTAAGATTTTCAAAGATGTTGTG 388

QY 478 GCCAGTATCTTACACAGAGAGATGAATGCAAGTGGGCTGGAAAGAGCATCTCTGAAAGA 537
DB |||||
DB 389 GCCAGTATCTTACACAGAGAGATGAATGCAAAATGGGCTGGAAAGATATCTCTGAAAGA 448
QY 538 ATGTGCTAAATTTTCATCAAGGTACTTAAAGCATATATATCAGACTCACTGTAGCCCTGTGG 597
DB |||||
DB 449 ATGTGCCAAATTTTCATCAAGGTCTCGGAGGCTTATATCAGACTCACTTGTATGSCCTGTGG 508
QY 598 AACGGGGCTTTTTCATCCAAATTTGCACTCAATGAAATTTGACATCATCTCTGAGGACAA 657
DB |||||
DB 509 AACTGGGGCTTTTCATCCAAATCTGACCTATATTTGAGTTGGACATCATCTCTGAGGACAA 568
QY 658 TATTTTAAAGCTGGGAAACTCACAATTTTGAATAACGGCTGGGAAGAGAGCCCTTATGATCC 717
DB |||||
DB 569 CATTTTAAAGCTGGAGGACTCACAATTTTGAATAACGGCTGGGAAGAGAGCCCTTATGATCC 628
QY 718 TAAGCTGTGACAGCATCCCTTTTAAATAGATGGAGAAATATATCTCTGAACTGACGTGA 777
DB |||||
DB 629 CAAACTACTGACTGCTCTCTTCTAATAGACGGTGAGTTGTACTCTGGAACCTGTGCGGA 688
QY 778 TTTTATGGGGCGAGACTTTTGTCTATCTTCCGAACCTCTTGGGCAACCAACCAATCAGGAC 837
DB |||||
DB 689 CTTTATGGGACGGGACTTTCGCTATCTTCAAGAACCTGGGGCAACCATCACCCCATCAGGAC 748
QY 838 AGAGCAGCATGATTCAGGTGGCTCAATGATCCAAAGTTTCATTTAGTGCCCACTCATCTC 897
DB |||||
DB 749 GGAGCAGCATGACTCCCGGTGGCTCAATGATCTTAGATTTCATCAGTGGCCATCTCATCCC 808
QY 898 AGAGATGACAAATCTCTGAAGATGACAAAGTATATCTTTTCTTCTGCGTGAATGCAATAGA 957
DB |||||
DB 809 AGAGATGACAAACCTCTGAAGATGACAAAGTATATTTTCTTCTCGAGAAATGCAATAGA 868
QY 958 TGGAGAACACTCTGGAAGAGTCTCAGCTAGATAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 1017
DB |||||
DB 869 CGGAGAACACTCTGGAAGAGGCACTCATGCTAGATAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 928
QY 1018 TGGAGGGCACAGAAAGTCTGTGTAATAATGGAACAACATTTCTCAAGCTCGTCTGATTTG 1077
DB |||||
DB 929 TGGTGGACACAGAAAGTCTTGTGAATAATGGAACAACATTTCTTAAAGCACGCTGATTTG 988
QY 1078 CTGAGTGGCGGTCGCAATGGCATTTGACATCTATTTGATGAACTGTGAGGATGATTTCT 1137
DB |||||
DB 989 CTCTGTGCGCGTCCCAATGGCAATTTGACACCAATTTTGTGAATTTGAGGATGATTTCT 1048
QY 1138 AATGAACCTTTAAAGATCTTAAATCCAGTTGTATGAGTGTGTATGAGTGTGTAGCTTCCAGTAA 1197
DB |||||
DB 1049 AATGAACCTTAAAGATCTTAAATCCAGTGTATGAGTGTGTATGAGTGTGTAGCTTCCAGTAA 1108
QY 1198 CATTTTCAAGGGATCAGCGGTGTGTATGTAATGATGATGATGTGAGAGGGTGTCTCT 1257
DB |||||
DB 1109 CATCTTTAAGGGATCTGCTGTGTGATGTACAGCATGATGATGTGTAAGAGGGTGTCTCT 1168
QY 1258 TGGTCCATATGCCCAAGGGATGAGCAACATCATCAATGGTGTGCTTATCAAGGAAGAGT 1317
DB |||||
DB 1169 TGGTCCATATGCTCACAGAGATGGTCCCACTATCAGTGGGTGCTTACCAAGGAAGAGT 1228
QY 1318 CCCCTATCCAGCGGACGAACTTCTCCAGCAAAACATTTGGTGGTGTGTGACTCTACAAA 1377
DB |||||
DB 1229 CCCCTATCCAGCGGACGAACTTGTCCAGTAAACATTTGGCGGATTTGATCTCCACAA 1288
QY 1378 GGACCTTCTCTGATGATTTATAACTTTTGAAGAGTCTCAGCCATGTACAAATCCAGT 1437
DB |||||
DB 1289 GGACCTTCTCTGATGATGTCAATCTTTTGAAGAGGATCATCCAGCCATGTACACCCAGT 1348
QY 1438 GTTTCCTATGAACAAATCGCCCAATAGTGATCAAAACCGAGTGTAAATATCAATTTACACA 1497
DB |||||
DB 1349 GTTTCCTATAAATAATCGCCGATCATGATCAAAACAGATGTAAATATCATGTTTACACA 1408
QY 1498 AATTGTGCTAGACCGAGTGGATGCAAGATGGAAGTGTATGATGATGATGATGATGATGATGATGAT 1557
DB |||||
DB 1409 AATTGTGCTAGACCGAGTGGATGCAAGATGGAAGTGTATGATGATGATGATGATGATGATGATGAT 1468

FASTSEQ for Windows Version 4.0
; SEQ ID NO 8352
; LENGTH: 2256
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(2256)
; OTHER INFORMATION: n = A,T,C or G
US-09-649-164-8352

Query Match 57.0%; Score 1544.8; DB 28; Length 2256;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1546; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY	59	CGACCCACCGCTCGGAGTAGTTGAGTCCGCTGTTCTCCATGTGACGACCTCTAT	118
DB	1	CGACCCACCGCTCGGAGTAGTTGTTGCTCGCTGTTCTCCATGTGACGACCTCTAT	60
QY	119	TTCCAGATTGTTGAACTTCTCTGGCCGACAAATACAGGAAGGAAGACTTAAAGCAGCAA	178
DB	61	TTCCAGATTGTTGAACTTCTCTGGCCGACAAATACAGGAAGGAAGACTTAAAGCAGCAA	120
QY	179	GGGACCTACAGCTCTGACGATGGCTGTTAACTAGGATTGTCTGTTTCTGGGA	238
DB	121	GGGACCTACAGCTCTGACGATGGCTGTTAACTAGGATTGTCTGTTTCTGGGA	180
QY	239	GTATTACTTACAGCAGCAACTATCAGATGGGAAGCAATGTGCCAAGCTGAAA	298
DB	181	GTATTACTTACAGCAGCAACTATCAGATGGGAAGCAATGTGCCAAGCTGAAA	240
QY	299	TTATCTTACAAAGAAATGTTGAATCCAAATGTGATCACTTTCAATGGCTTGGCCAA	358
DB	241	TTATCTTACAAAGAAATGTTGAATCCAAATGTGATCACTTTCAATGGCTTGGCCAA	300
QY	359	AGCTCCAGTTATCATACCTTCTTTGGATGAGAAACGGATGAGCTGTATGTTGAGCA	418
DB	301	AGCTCCAGTTATCATACCTTCTTTGGATGAGAAACGGATGAGCTGTATGTTGAGCA	360
QY	419	AGGATCATATTTTCACTGACCTGTTATATCAAGATTTCATAAGATTGTTGG	478
DB	361	AGGATCATATTTTCACTGACCTGTTATATCAAGATTTCATAAGATTGTTGG	420
QY	479	CCAGTATCTTACACAGAGAGATGAATGCAATGGCTGGAAAGACATCTCGAAAGAA	538
DB	421	CCAGTATCTTACACAGAGAGATGAATGCAATGGCTGGAAAGACATCTCGAAAGAA	480
QY	539	TGTGCTAATTCATCAAGTACTTAAAGCATATAATCAGACTCACTTGTACGCTGTGGA	598
DB	481	TGTGCTAATTCATCAAGTACTTAAAGCATATAATCAGACTCACTTGTACGCTGTGGA	540
QY	599	ACGGGGCTTTTCACTCCATTTGACCTACATTTGAATTTGGACATCATCTGAGGACAAT	658
DB	541	ACGGGGCTTTTCACTCCATTTGACCTACATTTGAATTTGGACATCATCTGAGGACAAT	600
QY	659	ATTTTAAAGCTGAGAACTCACATTTTGAACCGCGTGGGAAGATCCATATCACCT	718
DB	601	ATTTTAAAGCTGAGAACTCACATTTTGAACCGCGTGGGAAGATCCATATCACCT	660
QY	719	AAGCTGCTGACAGCATCCCTTTTAAATAGATGAGAAATTATCTCTGGAATGTCAGCTGAT	778
DB	661	AAGCTGCTGACAGCATCCCTTTTAAATAGATGAGAAATTATCTCTGGAATGTCAGCTGAT	720
QY	779	TTTATGGGCGAGACTTTGCTATCTCGAACTTTGGGCACCAACCCATCAGGACA	838
DB	721	TTTATGGGCGAGACTTTGCTATCTCGAACTTTGGGCACCAACCCATCAGGACA	780
QY	839	GAGCAGCATGATTCAGGTGGCTCAATGATCAAAAGTTTCATTAGTGGCCACCTCATCTCA	898
DB	781	GAGCAGCATGATTCAGGTGGCTCAATGATCAAAAGTTTCATTAGTGGCCACCTCATCTCA	840
QY	899	GAGAGTGACAATCTCTGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGAT	958

DB	841	GAGAGTGACAATCTCTGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGAT	900
QY	959	GGAGAACACTCTGGAAAGCTACTCAGCTAGAATAGTACAGATATGCAAGAACTGACTTT	1018
DB	901	GGAGAACACTCTGGAAAGCTACTCAGCTAGAATAGTACAGATATGCAAGAACTGACTTT	960
QY	1019	GGAGGACACAGAGTCTGGTGAATAAATGGACAACTTCTCAAAGCTCGTCTGATTTGC	1078
DB	961	GGAGGACACAGAGTCTGGTGAATAAATGGACAACTTCTCAAAGCTCGTCTGATTTGC	1020
QY	1079	TCAGTGCCAGTCCAAATGGCAATGACACTCATTTTGTATGAACTGCAGGATGATTTCTTA	1138
DB	1021	TCAGTGCCAGTCCAAATGGCAATGACACTCATTTTGTATGAACTGCAGGATGATTTCTTA	1080
QY	1139	ATGAACCTTTAAAGATCTTAAATCCAGTTCTATATGAGTGTATGAGTGTGAGTGTGAGT	1198
DB	1081	ATGAACCTTTAAAGATCTTAAATCCAGTTCTATATGAGTGTATGAGTGTGAGTGTGAGT	1140
QY	1199	ATTTTCAAGGATCAGCGCTGTGTATGTATAGCATGATGTGATGAGTGTGAGTGTGAGT	1258
DB	1141	ATTTTCAAGGATCAGCGCTGTGTATGTATAGCATGATGTGATGAGTGTGAGTGTGAGT	1200
QY	1259	GGTCCATATGCCCCACAGGATGGACCCAACTATCAATGGGTGCTTATCAAGGAAGATC	1318
DB	1201	GGTCCATATGCCCCACAGGATGGACCCAACTATCAATGGGTGCTTATCAAGGAAGATC	1260
QY	1319	CCCTATCCAGGCGCAGGAACTTGTCCAGCAAAACATTTGGTGTGTTTGTACTCTACAAAG	1378
DB	1261	CCCTATCCAGGCGCAGGAACTTGTCCAGCAAAACATTTGGTGTGTTTGTACTCTACAAAG	1320
QY	1379	GACCTTCTGATGATGTTTAAACCTTTTCAAGAAAGTATCCAGGCAATGATCAATCCAGTG	1438
DB	1321	GACCTTCTGATGATGTTTAAACCTTTTCAAGAAAGTATCCAGGCAATGATCAATCCAGTG	1380
QY	1439	TTTCTATGAAACAATCGCCCAATAGTATCAAAACGGATGTAATTTATCAATTTTACAAA	1498
DB	1381	TTTCTATGAAACAATCGCCCAATAGTATCAAAACGGATGTAATTTATCAATTTTACAAA	1440
QY	1499	ATTGTCGTAGACCGAGTGGATGCAAGAGATGCAAGATGATGATGTTTATCGGAACA	1558
DB	1441	ATTGTCGTAGACCGAGTGGATGCAAGAGATGCAAGATGATGATGTTTATCGGAACA	1500
QY	1559	GATGTTGGGACCGTCTTAAAGTAGTTTCAATTTCTTAAGGAGACTTGGTA	1608
DB	1501	GATGTTGGGACCGTCTTAAAGTAGTTTCAATTTCTTAAGGAGACTTGGTA	1550

Search completed: July 31, 2003, 19:21:14
Job time : 7231 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 13:13:31 ; Search time 326 Seconds
(without alignments)

4145.567 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709

Sequence: 1 aattctttatttcgatg.....aggettcttttcttaataacc 2709

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 424677 seqs, 249437934 residues

Total number of hits satisfying chosen parameters: 849354

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents_NA_New.*

- 1: /cgn2_6/ptodata/2/pna/pct_NEW_COMB.seq.*
- 2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
- 3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
- 4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
- 5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
- 6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
- 7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	647.4	23.9	2347	1	PCT-US03-09929-95
2	647.4	23.9	2347	1	PCT-US03-09929-97
3	647.4	23.9	2373	1	PCT-US03-09929-93
4	647.4	23.9	6474	1	PCT-US03-09929-91
5	556	20.5	5231	6	US-10-357-930-23105
6	556	20.5	5231	6	US-10-357-930-28973
7	407.8	15.1	1671	6	US-10-357-820-27
8	126.6	4.7	1438	1	PCT-US03-09929-25
9	126.6	4.7	2113	1	PCT-US03-09929-55
10	125	4.6	1327	1	PCT-US03-09929-21
11	125	4.6	1492	1	PCT-US03-09929-23
12	125	4.6	1878	1	PCT-US03-09929-33
13	125	4.6	1821	1	PCT-US03-09929-15
14	125	4.6	1948	1	PCT-US03-09929-49
15	125	4.6	2583	1	PCT-US03-09929-51
16	125	4.6	2634	1	PCT-US03-09929-53
17	125	4.6	2995	1	PCT-US03-09929-19
18	125	4.6	3055	1	PCT-US03-09929-27
19	125	4.6	3106	1	PCT-US03-09929-17
20	125	4.6	3165	1	PCT-US03-09929-47
21	125	4.6	3498	1	PCT-US03-09929-31
22	125	4.6	4250	1	PCT-US03-09929-13
23	123.4	4.6	2944	1	PCT-US03-09929-29
24	112.4	4.1	1814	1	PCT-US03-09929-9
25	110.8	4.1	1914	1	PCT-US03-09929-1
26	110.8	4.1	1914	1	PCT-US03-09929-7

27	110.8	4.1	2049	1	PCT-US03-09929-3	Sequence 3, Appli
28	109.2	4.0	2558	6	US-10-369-072-15	Sequence 15, Appli
29	109.2	4.0	3112	1	PCT-US03-09929-5	Sequence 5, Appli
30	109.2	4.0	3112	6	US-10-369-072-17	Sequence 17, Appli
31	106	3.9	2739	1	PCT-US03-09929-11	Sequence 11, Appli
32	106	3.9	3868	6	US-10-369-072-13	Sequence 13, Appli
33	94.4	3.5	1908	1	PCT-US03-09929-35	Sequence 35, Appli
34	92.8	3.4	3983	1	PCT-US03-09929-45	Sequence 45, Appli
35	90.6	3.3	1923	6	US-10-461-791-12	Sequence 12, Appli
36	90.6	3.3	1923	6	US-10-461-792-12	Sequence 12, Appli
37	90.6	3.3	3261	6	US-10-461-791-5	Sequence 5, Appli
38	90.6	3.3	3261	6	US-10-461-792-5	Sequence 5, Appli
39	90.6	3.3	3694	6	US-10-461-791-3	Sequence 3, Appli
40	90.6	3.3	3694	6	US-10-461-792-3	Sequence 3, Appli
41	87.4	3.2	4234	6	US-10-451-010-12	Sequence 12, Appli
42	76.4	2.8	8062	6	US-10-273-573-3841	Sequence 3841, Ap
43	63.8	2.4	2271	6	US-10-369-072-7	Sequence 7, Appli
44	63.8	2.4	2281	6	US-10-369-072-9	Sequence 9, Appli
45	63.8	2.4	3042	6	US-10-293-244-3941	Sequence 3941, Ap

ALIGNMENTS

RESULT 1

PCT-US03-09929-95

; Sequence 95, Application PC/TUS0309929

; GENERAL INFORMATION:

; APPLICANT: Curagen Corporation, et al

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD

; FILE REFERENCE: 21402-573B-061

; CURRENT APPLICATION NUMBER: PCT/US03/09929

; CURRENT FILING DATE: 2003-04-01

; PRIOR APPLICATION NUMBER: 60/368,996

; PRIOR FILING DATE: 2002-04-01

; PRIOR APPLICATION NUMBER: 60/369,980

; PRIOR FILING DATE: 2002-04-04

; PRIOR APPLICATION NUMBER: 60/370,381

; PRIOR FILING DATE: 2002-04-05

; PRIOR APPLICATION NUMBER: 60/370,969

; PRIOR FILING DATE: 2002-04-08

; PRIOR APPLICATION NUMBER: 60/371,002

; PRIOR FILING DATE: 2002-04-09

; PRIOR APPLICATION NUMBER: 60/372,002

; PRIOR FILING DATE: 2002-04-12

; PRIOR APPLICATION NUMBER: 60/384,297

; PRIOR FILING DATE: 2002-05-30

; PRIOR APPLICATION NUMBER: 60/386,816

; PRIOR FILING DATE: 2002-06-07

; PRIOR APPLICATION NUMBER: 60/389,123

; PRIOR FILING DATE: 2002-06-13

; PRIOR APPLICATION NUMBER: 60/402,207

; PRIOR FILING DATE: 2002-08-09

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 179

; SOFTWARE: Curaseqlist version 0.1

; SEQ ID NO 95

; LENGTH: 2347

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (14)..(2338)

PCT-US03-09929-95

Query Match 23.9%; Score 647.4; DB 1; Length 2347;

Best Local Similarity 57.6%; Pred. No. 5.5e-198;

Matches 1265; Conservative 0; Mismatches 906; Indels 24; Gaps 5;

QY 218 ATTGCTGCTTTCTTGGGAGTATTACTTACACAGACGACAACTATCAGATGGGAG 277

DB 35 ATCACCTTGCTCCTGTGGGGTTACTTACTGAGCTTTGGACAGAGGGTCATACAGCTGAT 94

Qy	278	AACAAATGTGCCAAGGCTGAAATATATCTTACAAAGAAATGTTGGSAATCCAAACAATGTGATC	337
Db	95	ACTACCACCCCGGTTACGGCTGTGCACATAAAGAGCTCTTTGAAATCTGGAACAGAACAATCA	154
Qy	338	ACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCATACCTTCCCTTTTGGATGAGGAACGG	397
Db	155	ATAATTTCAATAGCCCTTTTGGAAATTTCTTGATCTCCATACAATGCTGCTGGATGAATATCAA	214
Qy	398	AGTAGGCTGTATGTGGAGCAAAAGGATCACATATATTTTCATTCGACCTGGTTAAATATCAAG	457
Db	215	GAGAGGCTTCTGTTGGGAGCGGAGCACTTGTATATTTCCCTCAGCTGGAGAGAAATCAGT	274
Qy	458	GAT---TTTCAAAGAATGTGTGGCCAGTATCTTTACACCAAGAGAGATGAATGCAAGTGG	514
Db	275	GACGGCTATAAAGAGATACACTGGCCGAGTACAGCTCTAAAAAATGGAAGAAATGCTATAATG	334
Qy	515	GCTGGAAAAGACATCTTGAAAGAATGTGCTAAATTTTCATCAAGGTACTTAAAGGCATATAAT	574
Db	335	AAGGGAAAAGA---TGCGGGTGAATGTGCAAAATATATGTTGCGGTTTTTGCATCACTATAAC	397
Qy	575	CAGACTCACTTGTACGCTGTGGAAACGGGGCTTTTCATCCAAATTTGCACCTACATTGAA	634
Db	392	AGGACACACCTTCTGACCTGTGTACTGGAGCTTTTGNATCAGTTTGTGCTTCCATCAGA	451
Qy	635	ATTGGACATCATCTGTAGGACAAATATTTTTTAAGCTGGAGAACTCACATTTTGAAGAACGGC	694
Db	452	GTTCGATATCATTTGGAGGATCTCTGTTTTCACTGGAAATCACCCAGATCTGAGAGAGGA	511
Qy	695	CGTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGACATCCCTTTTAATAGATGGAGAA	754
Db	512	AGGGGACAGATGTCTTTTGTACCCAGCTCTCTCTTCATCTCCACTTTAAATGGTAGTGA	571
Qy	755	TTATACTCTGGAACTGACGCTGATTTTATGGCGCGAGACTTTTGCTATCTTCCGAACCTCTT	814
Db	572	TTGTTTGTGTGACTTACAGTGACTACTGGAGCAGAGAGCGTGCATCTTCCGCGAGCATG	631
Qy	815	GGGCACCAACCAATCAGACAGAGCAGCATGATTCAGAGTGGCTCAATGATCCAAAG	874
Db	632	GGGCGACTGGGCCATATCCGCACTGAGCATGACGATGAGCGTCTGTTGAAAGAACCAAAA	691
Qy	875	TTCAATTAGTGCACCTCATCTCAGAGAGTGACAAATCCTGAGATGACAAAGTATACCTTT	934
Db	692	TTTGTAGTTCATACATGATTCCTGCAATGAAGACAGAGATGACAAAGATATATTTTC	751
Qy	935	TTCTTCGTGAAAAATGCAATAGATGGAGAACACTCTGMAAAGCTACTCAGCTAGAGATA	994
Db	752	TTTTTTTACTGAAAGGCACTGGAGCAGAAAACAATGCTCAGGCAATTTACACCGGCTC	811
Qy	995	GGTCAGATATGCAAGAATGACTTTGGAGGGCAGCAAGAGTCTGTGTAATAAATGGACAACA	1054
Db	812	GGGCGACTCTGTGTGAATGATGTAGGAGGGCAGAGAAATCTGTGTGAATTAAGTGGAGCACT	871
Qy	1055	TTCCCTCAAAGCTCGTCTGAATTTGCTCAGTGCAGGTCGAAATGCGCATTGACACTATTTT	1114
Db	872	TTCCCTAAAGCGCAGACTCGTTTGTCTCAGTACAGGAATGAATGGAAATGACACATATTTT	931
Qy	1115	GATGAACTGACAGGATGTATTCCTAATGAACCTTTTAAAGATCCTAAAAATCCAGTTGTATAT	1174
Db	932	GATGAATTAGAGGACGTTTTTTTGTCTACTACAGAGATCATAAGAATCCAGTGATATTT	991
Qy	1175	GGAGTGTTTAGGACTCCAGGTAAACATTTTCAAGGGATCAGCCGCTGTGTATGTATAGCATG	1234
Db	992	GGACTCTTTAACTACACAGTAAATATTTTTCGAGGGCATGCTATATGTGCTTATCACAATG	1051
Qy	1235	AGTGATGTGAAGGGTGTCTTGTGTCATATGCCACAGGGATGGAACCAACTATCAA	1294
Db	1052	TCTAGCATTCGGGAGCGCTTCAACGGACCATATGCATATAGGAAGGACCTGATATCCAC	1111
Qy	1295	TGGTGCTTTATCAAGGAAGAGTCCCTCATCCAGCGCAGGAATCTTCTCCAGCAAAACA	1354
Db	1112	TGTCAGTCTATGAAGGAAGAANGTCCCTTATCCAAGGGCTGGTCTTGTGCGAGCAAAAGTA	1171
Qy	1355	TTTGGTG---GTTTGTGACTCTACAAAGGAGCCTTCTCGATGATGATTAATCACTTTGCAAGA	1411

Db	1172	AATGGAGGGAGATACGGAAACACCAAGGACTATCTCTGATGATGCCATCCGATTTGCAAGA	1231
Qy	1412	AGTCATCCAGCCATGTACAAATCAGTGTTCCTTATGAACAATCGCCCAATAGTGATCAAA	1471
Db	1232	AGTCATCCACTAATGTACAGGCCATAAACCCTGCCCATATAAAACCAATATTAGTAANA	1291
Qy	1472	ACGGATGTAAATATTCAATTTACACAATTTGCTAGACCGAGTGGATCGAGAAGATGGA	1531
Db	1292	ACAGATGGAAAAATATAACCTGAAACAATAAGCAGTAGATCGAGTGGAAAGCTGAGGATGGC	1351
Qy	1532	CAGTATGATGTTATGTTTATCGGAAACAGATGTTGGGACCGTTCCTTAAAGTAGTTCCAATT	1591
Db	1352	CAATATGACGCTCTGTTTATTGGGACAGATAATGGAAATTTGTGCTGAAAGTAATCACAATT	1411
Qy	1592	CCTAAGGAGACTTGGTATGATTTTAGAAGAGGTTCTCTCGAAGAAATGACAGTTCCTCG	1651
Db	1412	TACAAACCAAGAAATGGAAATCAATGGAAAGAAAGTAAATCTTAGAAGAACTTCAGATATTC	1471
Qy	1652	GAACCGACTGCTATTTTCAGCAATGGAGCTTTCACCTAAGCAGCAACAACATATATTTGGT	1711
Db	1472	GATCCAGTTCCTATTATTTCTATGGAGATTCTCTCAAAGCGCAACAGCTGTATATTGGA	1531
Qy	1712	TCAACGGCTGGGTTGCCAGCTCCCTTTACACCGGTGTGATATTTACCGGAAAGCGTGT	1771
Db	1532	TCTGCTCTGCTGGCTCAAGTCAGATTCATCACTGTGAATGTATGGAAGTGCCTGT	1591
Qy	1772	GCTGAGTGTTCCTCGCCCGAGACCTTACTGTGCTTGGGATGGTTCGCATGTTCTCGC	1831
Db	1592	GCTGACTGCTGCTGGCTCGAGACCCCTTACTGTGCTCGGATGGCATATCTTGCTCCCG	1651
Qy	1832	TATTTTCCCA-----TGCAAGAGACGCAAGACGCAAGATATAAGAAATGGA	1882
Db	1652	TATTTACCAACAGGCACACATGCAAAAGCGTTTCCGGAGACAAGATGTCGCACATGGA	1711
Qy	1883	GACCACTGACTCACTGTTTCAGACTTACACCATGATATCAACATGGCCACAGCCCTGAA	1942
Db	1712	AATGCACTCAGCAGTGCCTTTGGACAACTGTTGTTGGGATGCTTTGGATGAAGACTGAA	1771
Qy	1943	GAGAGAATCATCTATGCTGTAGAGAATAGTAGCACATTTTGGAAATGCACTCCGAAGTCG	2002
Db	1772	GAACATCTGGCTATTGGCATAGAAACAACAGTACTTTGCTGGATGTACCCCAAGATCT	1831
Qy	2003	CAGAGACGCTGCTCTATTGGCAATTTCCAGAGCGCAATGGAAGCGCAAAAGAAGATGC	2062
Db	1832	TTACAAGCGAAAGTTATCTGTTTGTACAGAAAGGACGTGAGACAAGAAAGAGGAGGTG	1891
Qy	2063	AGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGTAGTCTACAAAG	2122
Db	1892	AAGACAGATGACAGAGTGGTTAAGATGGACCTTGGTTTACTCTTCTTAAGGTTTACAAA	1951
Qy	2123	AAGGATTCAGGCAATTTACCTCTGCCATGCGGTGGAAACATGGGTTTCATACAAACTCTTCT	2182
Db	1952	TCAGATGCTGGGACCTATTTTGGCCAGACAGTAGAGCATAGCTTTGTCCATACGGTCCGT	2011
Qy	2183	AAGGTAAACCTCGAAGTCAATGACACAGAGCAATTTGGAGAAGACTTCTTCATTAAGAGATGAT	2242
Db	2012	AAAATCACCTTGGAGGTAGTGGAAAGGAGAAAGTCGAGGATATGTTTAAACAAGGACGAT	2071
Qy	2243	GATGGAGATGGCTCTTAAGACCAAGAAATGTC-----CAATAGCATGACACCTAGCCAG	2296
Db	2072	GAGGAGGACAGGCAATCACAGGATGCTCTGCTCTCAGAGTAGCATCTCGCAGGAGCA	2131
Qy	2297	AAGGTCTGGTACAGAGACTTCATGCACTCATCAACCAACCCCAATCTCAACACGATGGAT	2356
Db	2132	AAACCATGGTACAAGGAAATCTTTCGACGTGATCGGTTATAGCACTTCAGAGAGTGGAA	2191
Qy	2357	GAGTTCGTGGAACAAGTTTGGAAAAAGGGACCGAAA	2391
Db	2192	GAATATCGCAGAAAGTATGTTGCAAGATAGAAA	2226

RESULT 2

Query Match	23.9%;	Score 647.4;	DB 1;	Length 2347;
Best Local Similarity	57.6%;	Pred. No. 5.5e-198;		
Matches 1265;	Conservative	0;	Mismatches 96;	Indels 24;
Gaps	5			
QY	218	ATTGTCTGTCTTTCTGGGGAGTATTACTTACAGCAGAGCAAACTATCAGAAATGGGAG	277	
DB	35	ATCACCTTGGCTCTGGGGTTACTTACTGGAGCTTTGGACAGAGGTATACAGCTGAT	94	
QY	278	AACATGTGCCAAGCTGAAATTATCTTACAAGAAATTTTGGAAATCCCACAATGTGATC	337	
DB	95	ACTACCCACCCCGGTACGCGTGTGCATAAAGAGCTCTTGAATCTGAACAGAACATCA	154	
QY	338	ACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCATACCTTCCTTTTGGATGAGGACGG	397	
DB	155	ATATTTCATAGCCCTTTTGGATTTCTTGATCTCTACAATGCTGCTGGATGAATATCAA	214	
QY	398	AGTAGCTGTATGTTGGAGCAAGGATCACATATTTTTCATTGCACCTGGTTAATATCAAG	457	
DB	215	GAGAGGCTCTTGGTGGAGGACGGACCTTGATATATTCCTCAGCTTGGAGAGAAATCA	274	
QY	458	GAT---TTTCAAAAGATTGTGCGCAGTATCTTACACAGAGAGATGAATGCAATGG	514	
DB	275	GACGGCTATTAAGAGTACACTGCGCGAGTACAGCTCTAAAATGGAAGAATGCATATG	334	
QY	515	GCTGAAAAGACATCCTGAAAAGATGTGCTAAATTCATCAAGGTACTTAAAGGCATATA	574	
DB	335	AAGGGAAAAGA---TGCGGGTGAATGTGCAAAATATGTTCGGGTTTGGCATCACTATA	391	
QY	575	CAGACTCATCTGTACGCTCTGTGGACGGGGCTTTTCATCCATTTTGCACCTTACATT	634	
DB	392	AGACACACCTTCTGACCTGTGGTACTGGAGCTTTTGTATCCAGTTTGTGCTTTCAT	451	
QY	635	ATTGGACATCATCTCCAGGACAAATATTTTTAAAGCTGGAGAACTCACATTTTGAAAC	694	

Db 821 GGGGACTCTGTGTAATGATGTAGGAGGCGCAGAGAACTACTGGTGAATAGTGGAGCACT 880
Qy 1055 TTCTCTAAAGCTGTGCTGATTTGCTGAGTCCAGGTCCAAATGGCAATGACACTCATTTT 1114
Db 881 TTCTCTAAAGCGAGACTCGTTGCTCAGTACCAGGAATGAATGGAATTCACACATATTTT 940
Qy 1115 GATGAAGTCCAGGATGATTTCTTAATGAATTTAAAGATCCCTAAATCCAGTGTGTAT 1174
Db 941 GATGAATTAGAGGAGCTTTTTTGTCTACCTACAGAGATCATGAAGATCCAGTGTATTT 1000
Qy 1175 GGAGTGTTTACGACTTCCAGTAACATTTTCAAGGATCAGCCGTGTGTATGTATAGCATG 1234
Db 1001 GGAATCTTTAAGTACAGTACAGTATATTTTTCAGGCGATGCTATATGTCTATCAGATG 1060
Qy 1235 AGTATGTGAGAGGCTGCTTCCATATGCCCCAGGAGTGGACCCCACTATCAA 1294
Db 1061 TCTAGCATTCGGGCGACCTTCAACGGACCATATGACATAAGGAAGGACCTGAATACCAC 1120
Qy 1295 TGGGTGCTTATCAAGGAAGTCCCTTATCCAGGCGCAGGAACCTTGTCCAGGAAACA 1354
Db 1121 TGGTCAGTCTATGAAGGAAGTCCCTTATCCAGGCGCTGTTCTTGTGCCAGCAAGTA 1180
Qy 1355 TTGTGTG---GTTTGTACTCTACAAAGGACCTTCTGTGATGATGTTATTAACCTTTCCAAGA 1411
Db 1181 ATGGAGGAGATACGGAACACCAAGGACTATCTGTATGATGCCATCGATTCGAAGA 1240
Qy 1412 AGTCATCCAGCATGTACAAATCCAGTGTTCCTATGAACAAATCGCCCAATAGTATCAA 1471
Db 1241 AGTCATCCACTAATGTACAGGCCATAAAACCTGCCCATATAAAACCAATATAGTAAA 1300
Qy 1472 ACGGATGTAAATATCAATTTACAAATTTGCTGAGACCGAGTGGATGCAGAAATGGA 1531
Db 1301 ACAGATGGAAATATAAATCTGAACAAATAGCAGTAGATCGAGTGGAAAGCTGAGATGCG 1360
Qy 1532 CAGTATGATGTTATGTTATCGGAACAGATGTTGGACCGTCTTAAAGTAGTTCATTT 1591
Db 1361 CAATATGACGCTGTTTATGCGACAGATGAATGGAATGTGCTGAAAGTATCAAAAT 1420
Qy 1592 CTAAGGAGACTTGTGTATGATTTAGAAGAGTTCCTGCTGGAAGAAATGACAGTTCCTCGG 1651
Db 1421 TACAACCAAGAAATGGAATCAATGGAAGAAATTTCTAGAAGAACTTCAGATATTCAG 1480
Qy 1652 GAACGACTGTATTTACGAATGAGCTTTCCACTAAGACGCAACAATATATATTTGGT 1711
Db 1481 GATCAGTTCCTATATTTCTATGAGATTTCTTCAAGCGGCAACAGCTGTATATGGA 1540
Qy 1712 TCAACGCTGGGTTGCGCCAGCTCCCTTTACACCGGTGTGATTTTACGGGAAGCGTGT 1771
Db 1541 TCTGCTTCTGCTGGCTCAAGTCAGATTCATCACTGTGACATGTATGGAAGTCTTGT 1600
Qy 1772 GCTGAGTGTGCTCGCCGACACCTTACTGTGCTGGATGGTTCCTGCATGTTCTCGC 1831
Db 1601 GCTGACTGCTGCTGCTGAGACCTTACTGTGCTGGATGGATATCTGCTCCCGG 1660
Qy 1832 TATTTTCCAC-----TGCAAAGACGCAACAGACGCAAGATATAGAATGGA 1882
Db 1661 TATTACCCCAACAGGCACATGCAAAAAGCGCTTTCCGAGACAGATGTTCCGACATGGA 1720
Qy 1883 GACCCACTGACTCACTGTTCAGCTTACACATGATATCACTATGGCCAGCCCTGAA 1942
Db 1721 ATGAGCTCAGCAGTCTTTGGAACAACAGTTTGTGGGATGCTTTGGATAAGATGAA 1780
Qy 1943 GAGAGATCATCTATGCTGTAGAAATAGTACACATTTTGGAAATGAGTCCGAAGTCG 2002
Db 1781 GAACATCTGGCTTATGGCATAGAAACACAGTACTTTGCTGGATGTACCCCGAGTCT 1840
Qy 2003 CAGAGAGCGCTGCTTATTTGGCAATTCAGAGCGGAAATGAAGCGCAAAAAGAGATC 2062
Db 1841 TTACAAGCGAAAGTTATCTGTTTCTGACAGAAAGGACGTGACACAAGAAAGAGGAGGTG 1900
Qy 2063 AGAGTGGATGATCATATCAGGACAGATCAAGSCCTTCTGCTACGTAGTCTACAAAG 2122
Db 1901 AAGACAGATGACAGAGTGGTTAAGATGGACCTTGTGTTTACTCTTCTTAAGGTTACACAA 1960

Qy 2123 AAGGATTCAGCAATATACCTCTGCCATCGGTGGAAATGGTTTCATACAACTCTTCTT 2182
Db 1961 TCAGATGCTGGGACCTATTTTTCAGACAGCTAGAGCATAGCTTTGTCCATACGCTCCGT 2020
Qy 2183 AAGGTAACCTTGGAGTCAATTCACACAGCATTTGGAGAACTTCTTCATAAAGATGAT 2242
Db 2021 AAAATCACTTTGGAGGTAGTGGAAAGAGGAAAGTTCGAGGATATGTTTAAACAAGGACGAT 2080
Qy 2243 GATGAGATGCTCTTAAGACCAAAAGAAATGTC-----CAATAGCATGACACCTAGCCAG 2296
Db 2081 GAGGAGACAGGCATCACAGATGCCTTGTCTGCTCAGATAGCATCTCGCAGGAGCA 2140
Qy 2297 AAGTCTGTGATCAGAGACTTTCATCAGCTCATCAACCCCAATCTCAACACATGATGAT 2356
Db 2141 AAACCATGTTCAAGGAATTTCTGAGCTGATCGTTTATAGCAACTTCCAGAGAGTGGAA 2200
Qy 2357 GAGTCTCTGAACAAGTTTGGAAAAGGACCGAAA 2391
Db 2201 GAATACTCGGAGAAGTATGGTGCACAGATAGAAA 2235

RESULT 4

PCT-US03-09929-91
; Sequence 91, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraSequid version 0.1
; SEQ ID NO 91
; LENGTH: 6474
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (467)..(2791)
PCT-US03-09929-91

Query Match 23.9%; Score 647.4; DB 1; Length 6474;
Best Local Similarity 57.6%; Pred. No. 9.2e-198;
Matches 1265; Conservative 0; Mismatches 906; Indels 24; Gaps 5;
Qy 218 ATTCTGTCTCTTTCTGGGAGTATTACTTACAGCAAGAGCAAACTATCAGAATGGGAAG 277
Db 488 ATACCTTGTCTCTGTTGGGTACTTACTGAGCTTTGGACAGGAGGTCTACAGCTGAT 547
Qy 278 AACATGTGCAAGCGCTGAAATTTATCTCAAAAGAAATGTTGGAATCCAAACATGTGATC 337

Db 548 ACTACCCACCCCGGTTACGCTGTGCATATAAGAGCTCTTGAATCTGAACAGACATCA 607
Qy 338 ACTTTCAATGGCTTGGCCACAGCTCCAGTTATATACCTTCTCTTTGGATGAGGACGG 397
Db 608 ATATTTTCATAGCCCTTTTGGATTTCTTGATCTCCATACAAATGCTGCTGGATGAATCAA 667
Qy 398 AGTAGGCTGTATGTTGGAGCAAGGATCAATATTTTCATTCGACCTGGTGTAAATCAAG 457
Db 668 GAGAGGCTCTTGGTGGAGGCGAGGACCTTGATATCTCCCTCAGCTTGGAGAGATCAGT 727
Qy 458 GAT-...TTTCAAAGATTTGTGGCCAGTATCTTACACAGAGAGATGAATGCAAGTGG 514
Db 728 GACGGCTATAAAGAGATACACTGGCCGAGTACAGCTCTTAAATGGAAGATGCAATATG 787
Qy 515 GCTGGAAGAGACATCTGGAAGATGTGCTAATTTTCATCAAGGTACTTAAAGCCATATAT 574
Db 788 AAGGGAAGAG-...TGCGGGTGAATGTGCAATATATGTTCCGGTTTTGCATCACTATAC 844
Qy 575 CAGACTCACTTGTACGCTGTGGAACGGGGCTTTTTCATCCAAATTTGCACCTTACATGAA 634
Db 845 AGGACACACTTCTGACCTGTGTACTGGAGCTTTTGATCCAGTTTGTGCTTTCATCAGA 904
Qy 635 ATTGACATCATCTGAGGACAAATATTTTAAAGTGGAGAACTCAATTTTGAAGAACGGC 694
Db 905 GTTGGATATCATTTGGAGGATCTCTGTTTCACTTGAATCACTCCAGATCTGAGAGAGA 964
Qy 695 CGTGGAGAGTCCATATACCTTAAGCTGTGACAGCATCTCTTTTAAATAGATGAGAA 754
Db 965 AGGGGAGATGTCTTTTACCCAGCTCTCTCTTCACTTCACTTTTAAATGGTAGTGA 1024
Qy 755 TTATCTCTGGAAGTGCAGCTGATTTTATGGGCGAGACTTGTCTATCTTCCGAACTCTT 814
Db 1025 TTGTTTGTCTGACTCTACAGTGACTACTGAGCAGAGAGCTGCGATCTTCCGAGCATG 1084
Qy 815 GGGCACCACCAATCAAGAGCAGAGCAGATGATTCAGGTGCTCAATGATCAAAG 874
Db 1085 GGGCGACTGGCCCATATCCGCACTGAGCATGACGATGAGCGTCTGTGAAAGAACAAA 1144
Qy 875 TTCAATAGTCCACCTCATCTCAGAGTGAATCTCTGAGATGACATCTCTGAGATGACAAAGTATACTTT 934
Db 1145 TTGTAGGTTTCAATGATTTCTGACATGAAGAGCAGAGATGACAAAGTATATTTTC 1204
Qy 935 TTCTTCGTTGAAATGCAATAGATGAGAACTCTGGAAGAGTACTCACTGCTAGAATA 994
Db 1205 TTCTTTACTGAGAGGCACTGGAGCAGAAACATGCTCAGCAATTTACACAGGCTC 1264
Qy 995 GGTGAGATGCAAGATGACTTTGGAGGCGACAGAGTCTGGTGAATTAATGGAACA 1054
Db 1265 GGGCGACTCTGTGTAATGATGTAGGAGGCGAGAAATACCTGGTGAATTAAGTGAGCACT 1324
Qy 1055 TTCTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCAGGTCCTCAATGCGCATTTT 1114
Db 1325 TTCTTAAAGCGAGACTCGTTTGTCTCAGTACCAGGAATGAATGGAATTTGACATATTT 1384
Qy 1115 GATGAATCTCAGGATGATTTCTCAATGAATTTTAAAGATCTTAAAGATCTTCAATTTAT 1174
Db 1385 GATGAATTAGAGGACGTTTTTTTGTCTACCTACAGAGATCAATGAATCCAGTATTT 1444
Qy 1175 GGAGTGTTCAGCTTCCAGTAACTTTTCAAGGAGATCAGCGTGTGTATGATAGCATG 1234
Db 1445 GGACTCTTTAACTACCACTACCACTAATATTTTTCGAGGCGATGCTATATGCTCTATCAGATG 1504
Qy 1235 AGTGATGTGAGAGGCTGTTCTTGTCTCATATGCCACAGGATGGACCACTATCAA 1294
Db 1505 TCTAGCATTCGGGCGCTTCAACGGACCATATGACATAAGGAAGGACCTTGAATACCAC 1564
Qy 1295 TGGGTGCTTTTCAAGAGAGTCCCTTATCCAGCGCCAGGAACTTTGTCGCCAGCAAAACA 1354
Db 1565 TGTGAGTCTAAGAGGAAAGTCCCTTATCCAGGCTGTTCTTGTGTCAGCAAGATA 1624
Qy 1355 TTGTGTG-...GTTTGTACTCAAGAGGACCTTCTCTGATGATGTTTAACTTTGCAAGA 1411
Db 1625 AATGGAGGAGATACGGAAACCAAGGACTATCTGTATGATGCTCATCGATTTGCAAGA 1684

Qy 1412 AGTCATCCAGCATGTACAATCCAGTGTCTCTATGAACAATCGCCCAATAGTGTATCAA 1471
Db 1685 AGTCATCCACTAATGTACAGGCCATAAACTGCCCATAAACCAATATTTGGTAAA 1744
Qy 1472 ACGGATGPAATATTAATTTTACAAATTTCTGTAGACCGAGTGGATGCAAGATGGA 1531
Db 1745 ACAGATGGAATAATATACTGAAACNAATAGCAGTAGATCGAGTGAAGCTGAGGATGCG 1804
Qy 1532 CAGTATGATGTTATTTTATCGGAACAGATGTTGGGACCGTCTTAAAGTAGTTCATTT 1591
Db 1805 CAATATGACGCTCTGTTTATTTGGGACAGATAATGGAATTTGCTGAAAGTAAATCAAT 1864
Qy 1592 CCTAAGGAGACTTGTATGATTTTGAAGAGGTTCTGTGGAAGAATGACAGTCTTTTGG 1651
Db 1865 TACAACCCAGAAATGGAATCAATGGAAGAGTAATCTTAGAAGAACTTCAGATATTCAG 1924
Qy 1652 GAACCGACTGCTATTTTTCAGCAATGGAGCTTTCCATTAAGCAGCAACAACTATATTTGGT 1711
Db 1925 GATCCAGTTCTATTTATTTCTATGGAGATTTCTTCAAACCGCAACAGCTGTATATTGGA 1984
Qy 1712 TCAAGGCTGGGGTTGCCAGCTCCCTTTACACGGTGTGATATTTACGGGAAGCGTGT 1771
Db 1985 TCTGCTTCTGCTGTGGCTCAAGTCAGATTCATCACTGTGACATGTATGGAAGTCTTGT 2044
Qy 1772 GCTGAGTGTTCCTCGCCGAGAGCTTACTGTCTGTGGGATGGTCTGCAATGTTCTCGC 1831
Db 2045 GCTGACTGCTGCTGCTGCGAGACCTTACTGTGCTGGGATGGCATATCTGCTCCCGG 2104
Qy 1832 TATTTTCCAC-...TSCAAAGAGACGCAAGACAGCAAGATATTAAGAAATGGA 1882
Db 2105 TATTACCCAAACAGGCACACATGCAAAAAGGCGTTTCCGAGACAAAGATGTTTCGACATGGA 2164
Qy 1883 GACCACTGACTCTGTTTACAGCTTACACCATGATATACCATGGGCCACAGCCCTGAA 1942
Db 2165 AATCAGCTCAGCAGTGTCTTTGGAACAAGTGTGTTGGGATGCTTTGGATAAGACTGAA 2224
Qy 1943 GAGAGATCATCTATGTTGTAGAGAAATAGTAGACATTTTGGGAATGAGTCCGAAGTCG 2002
Db 2225 GAACTCTGGCTTATGGCATAGAGAAACACAGTACTTTGCTGGATGTATCCCCAGCATCT 2284
Qy 2003 CAGAGAGCGTGTCTTATTTGGCAATTCAGAGGCGAAATGAAGAGCGGAAAAAGAGATC 2062
Db 2285 TTACAAGCGAAAGTTATCTGTTTGTACAGAAAGGACGCTGAGACAAAGAAAGAGAGGTG 2344
Qy 2063 AGAGTGTATGATCATATCATCAGCAGATCAAGGCTTCTGCTACGTAGTCTACAAAG 2122
Db 2345 AAGCAGATGACAGAGTGTGTAAGATGACCTTGGTTTACTCTTCTTAAAGGTTACACAAA 2404
Qy 2123 AAGGATTCAGGCAATTTACTCTGCGATGCGGTGGAACATGGGTTTCATACAACTCTTCTT 2182
Db 2405 TCAGATGCTGGACCTATTTTTCAGACAGTATGAGCATAGCTTTTGTCCATACGCTCGT 2464
Qy 2183 AAGGTAAACCTTGGAGTCAITGACACAGAGCATTTGGAAGAACTTTCTTCAATAAGATGAT 2242
Db 2465 AAAATCACTTTGGAGGTAGTGAAGAGAGAGAAAGTCGAGGATATGTTTAAACAGACGAT 2524
Qy 2243 GATGAGATGCTCTAAGACCAAGAAATGTC-...CAATAGCATGACACCTAGCCAG 2296
Db 2525 GAGGAGGACAGGCATCAAGGATCCCTTGTCTCTCAGAGTAGCATCTCCGAGGAGCA 2584
Qy 2297 AAGGTCTGGTACAGAGACTTTCATGCTCATCAACCAACCCCAATCTCAACAGATGAT 2356
Db 2585 AAACCATGTTACAGGAATTTCTGACGTGATCGTTATAGCAACTTCCAGAGTGA 2644
Qy 2357 GAGTCTGTGAACAAAGTTTGGAAAGGACCGAAA 2391
Db 2645 GAACTCTGCGAGAAAGTATGTTGTCACAGATAGAAA 2679

RESULT 5
US-10-357-930-23105
; Sequence 23105, Application US/10357930

GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Delege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23105
; LENGTH: 5231
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 5231
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-23105

Query Match 20.5%; Score 556; DB 6; Length 5231;
Best Local Similarity 57.6%; Pred. No. 2.5e-168;
Matches 1117; Conservative 0; Mismatches 800; Indels 21; Gaps 6;

QY	288	CAAGCTGAAATATCTACAAAGAAATGTTGGATCCCAACATGATGATCATTCAATG	347
Db	642	CAAGAGTTTATTAATTTGATGAACTTCGAGAAACCAAGACCTCTGAAATCTCAGCC	701
QY	348	GCTTGGCCCAACAGCTCCAGTTATACCTCTCTTTGGATGAGGAACGGAGTAGGCTGT	407
Db	702	TTTCCCAACATCTTTAGACTACAGGATTTATTAATGATGAAATCAGGACCGATAT	761
QY	408	ATGTTGGAGCAAGGATCACATATTTTCATTCGACCTGGTTAAATCA---AGGATTTTC	464
Db	762	ATGTGGGAAGCAAGATCACATTTCTTCCCTGAAATATTAACAAATATAAGTCAAGAAGCTT	821
QY	465	AAAAGATTGTGGCCAGTATCTTACACAGAGAGATCAATGCAAGTGGGCTGGAAG	524
Db	822	TGAGTGTCTTCTGGCCAGATCTCAATCAAAAGTTGAAGATGCAAAATGGCTGGCAAG	881
QY	525	ACATCTCTGAAAGAAATGTCTAATTTTCATCAAGGTACTTAAGGCATATATCAAGCTCACT	584
Db	882	ATCCACACACAGCTGTGGAACTTTGTCGGTGAATTCAGACTTTCAATCGCACATTT	941
QY	585	TGTACGCTGTGGAACGGGGCTTTTCATTCACAAATTTGACCTTACATTAATAATGGAATC	644
Db	942	TGTATGTCTGTGGAGTGGCGCTTTTCAGTCTCTGTCTGTACTTACTTGAACAGAGGAGGA	1001
QY	645	ATCCTGAGGACAAATTTTAAAGTGGAGATCAATTTGAAACCGCGTGGGAAGA	704
Db	1002	GATCAGAGGACCAAGTTTTCATGATG---ACTCCAAGTGTGAATCTGGAAGAGGACGCT	1058
QY	705	GTCCATATGACCTCAAGCTGTGACAGCATCCCTTTTAATAGATGGAAATATATCTCTG	764
Db	1059	GCTCTTTCACACCCCAAGTGACAGGCTGTCTGTTATGATCATGATGAGGAGCTTTCTCTG	1118
QY	765	GAACTGACGTGATTTTATGGGGGAGAGCTTTTGCTATCTTCCGAATCTCTGGGCCACC	824

Db	1119	GAATGTATATAGATTTTCATGGGACAGATGCTGCTATTTTTCGAAGTTTAAACCAAGAGA	1178
QY	825	ACCCAATCAGACAGACAGCATGATTCAGCTGGCTCAATGATCCAAAGTTCATTAGTG	884
Db	1179	ATGCGGTGAGAACTGATCAACATATTTCCAATGGCTAAGTGAACTATGTTGTAGATG	1238
QY	885	CCCACCTCATCTCAGAGAGTGACAAATCCTGAAGATGACAAAGATATCTTTTTCCTCGTG	944
Db	1239	CACATGTATCCAGATGGTACTGATCCAAATGATGCTAAGGTGATCTTCTTCTTCAAAG	1298
QY	945	AAATGCAATAGATGGAGAACACTCTGGAAGCTTACTCAGCTAGATAGTGCAGATAT	1004
Db	1299	AAAACTGACTGACAAATAACAGAGACGAAACAGATTCATTTCCATGATGCTCGAATAT	1358
QY	1005	GCAAGATGATTTGGAGGGACAGAACTCTGTTGAATATAATGGAACAACATTCCTCAAG	1064
Db	1359	GTCTTAATGACACTGTTGGACTGCTGAGCTTGTCAACAAGTGGACCACTTCTTAAGG	1418
QY	1065	CTCGTCTGATTTGCTCAGTCCAGGTCCAAATGGCATTTGACACTCATTTTGAAGACTGC	1124
Db	1419	CGAGCTGGTGTGCTCGGTAAACAGATGAAGACGGCCAGAAACACACTTTGATGAATAG	1478
QY	1125	AGGATGATTTCTTAATGAACTTTAAAGATCCTAAAAATCCAGTTGTATATGAGGTGTTA	1184
Db	1479	AGGATGTTTCTGCTGGAAACTGATAACCCGAGGACCAACACTAGTGTATGGCATTTTA	1538
QY	1185	CGACTTCCAGTAAACATTTTCAAGGATCAGCGCTGTGTATGTATATGATGATGATGTA	1244
Db	1539	CAACATCAAGCTCAGTTTCAAGATCAGCGCTGTGTGTGTATCAATTTCTGATATAC	1598
QY	1245	GAAGGTTTCTTCTTGGTCCATATGCCACAGGGATGGACCAACTATCAATGGGTGCCTT	1304
Db	1599	AGACTGTGTTTAAATGGGCTTTTGGCCACAAAGAGGGCCCAATCATCAGCTGATTTCT	1658
QY	1305	ATCAAGAGAGTCCCTTATCCAGCGGACGAACTTGTGCCAGCAAAACATTT---GGTG	1361
Db	1659	ATCAGGCGAATTTCCATATCTCGCCCTGGAACTTGTCCAGGAGGACATTTTACACCA	1718
QY	1362	GTTTGTACTCTACAAAGGACCTCTCTGATGATGTTTATAACCTTTTCAAGAGATCATCCAG	1421
Db	1719	ATATCGGNACACAGAGGATTTCCAGATGATGTTGTCACTTTTATTCGGAACCACTCTC	1778
QY	1422	CCATGTACAAATCCAGTGTTCCTATGAACAATCGCCCAATAGTATGATCAAAAACGGATGA	1481
Db	1779	TCATGTACAAATTCATCTACCAATCCACAAAGGCTTTGATTTGCTATTTGGCACTG	1838
QY	1482	ATTATCAATTTACAAATTTGCTGACAGCTGGATGAGAGATGGACAGTATGATG	1541
Db	1839	ACTACAAGTATACAAAGATAGCTGTGGATCGAGTGAACGCTGCTGATGGGAGATACCAATG	1898
QY	1542	TTATGTTTATCGGAACAGATGTTGGGACCGTCTTAAAGTAGTTCATTTCTTAAGGAGA	1601
Db	1899	TCCTGTTTCTCGGAACAGATCGGGGTCTGTGCAAAAAGTGTGTTGTTCTTCTACTAACA	1958
QY	1602	CTTGATGTATTTAGAAGAGGTTCTGCTGGAAGAAATGACAGTTTTTCGGGAACCGACTG	1661
Db	1959	---ACTGCTGAGTGGGAGCTCATCTGAGGAGCTGGAAGTCTTTTAAAGATCATGCTC	2015
QY	1662	CTATTTGAGCAATGGAGCTTTCCATAGAGAGCAACATATATATTTGGTTCAAGCGCTG	1721
Db	2016	CTATAACAACAATGAAATTTTTCATCTAAAAAGCAACAGTTGTATGTAGTTCCTCAATGAAG	2075
QY	1722	GGTTTGGCCAGCTCCCTTTTACACCGTGTGATATTTAGGGAAGACGCTGCTGAGTGT	1781
Db	2076	GGGTTTCCAGGATATCTGACCGCTGCCACATCTATGGTACAGCTTGTGCTGACTGCT	2135
QY	1782	GCCTCGCCGAGACCTTTACTGCTTTGGATGTTTCTGCAATGTTCTGCTTATTTTCCCA	1841
Db	2136	GCCTGGCGGGACCTTTATTTGGCTGGATGGCCATCTCTGTTCCAGATTTCTACCCAA	2195
QY	1842	CTGCAAGAGACGCAAGAGCAGCAATATTAAGAAATGGAGACCCACTGACTGCTGCTT	1901

Db 2196 CTGGAAACGAGGCGCGAAGACAGATGTGAGACATGGAACCCACTGACTCAATGCA 2255
Qy 1902 CAGACTTACACCATGATATACCATGGCCACAGCCCTGAAGAGAGAAATCATCTATGGTG 1961
Db 2256 GAGGATTT-----TAATCTAAAGCATACAGAAATGCAGCTGAAATTTGGCAGTATGGAG 2309
Qy 1962 TAGAATATAGTAGCACAATTTTGGAAATGCGAAGTCGAGAGAGCGCTGGTCTATT 2021
Db 2310 TAAAAAATAACCACTTTCTGGAGTGCGCCCAAGTCTCCGACGACATCTATCAAGT 2369
Qy 2022 GCGAATTCAGAGGCGGAATGAAGAGCGGAAGAGAGATGAGATGATATATCA 2081
Db 2370 GCGCTGTATACAGA---AAGACAAGACAGGAGGAAGAGGTTAAGCTGAAATGAACGAATA 2426
Qy 2082 TCAGACAGACATCAAGSCCTTCTGTACGTAGTCTCAACAGAGAGATTCCAGGCAATTACC 2141
Db 2427 TAGCCACTTCACAGGACTCTCGATCCGCTCTGTTTCAAGGTTCTGACCAAGGACTTTATC 2486
Qy 2142 TCTGCCATCGGTGGAAACATGGGTTTCATACAAACTCTTCTTAAGGTAAACCTCGGAAGTCA 2201
Db 2487 ACTGCATTCTACAGAAATAGTTTCAACGACACCATACCCAGATCAACTTCAAGTTT 2546
Qy 2202 TTGACACAGAGCAATTGG 2219
Db 2547 TAGATTCAAAATGGTGG 2564

RESULT 6

US-10-357-930-28973
; Sequence 28973, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28973
; LENGTH: 5231
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..2..5231
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-28973

Query Match 20.5%; Score 556; DB 6; Length 5231;
Best Local Similarity 57.6%; Pred. No. 2.5e-168;
Matches 1117; Conservative 0; Mismatches 800; Indels 21; Gaps 6;
Qy 288 CAAGCTGAAATATCTTCAAGAAATGTTGGAATCCAAATGTGATCACTTTCAATG 347
Db 642 CAAGAGTTTATTAAACATTTGATGAATTCGAGAAACCAAGACCTCTGAATACTTCAGCC 701

Qy 348 GCTTGGCCAAACAGCTCCAGTTTATCATACCTTCTCTTTGGATGAGGAACGAGTAGCTGT 407
Db 702 TTTCCACCATCTTTAGACTACAGGATTTTATTAATGATGAAGATCAGGACCGGATAT 761
Qy 408 ATGTTGGAGCAAGGATCACATATTTTCATTCGACCTGGTTAATATCA---AGGATTTTC 464
Db 762 ATGTGGGAAGCAAGATCACATCTTTCCCTGATATTAACAATATAAGTCAAGAAGCTT.821
Qy 465 AAAGATTTGTGGCCAGTATCTTACACAGAGAGATGAATGCAAGTGGGCTGGAAAG 524
Db 822 TGAGTGTTTTCTGGCCAGCATCTACAATCAAGTTTGAAGAATGCAAAATGGCTGGCAAG 881
Qy 525 ACATCTCTGAAAGATGTCTAATTTTCATCAAGGTACTTTAAGGCATATATATCAGACTCACT 584
Db 882 ATCCACACACGGCTGTGGGAACCTTGTCCGTGTAATTCAGACTTTCATCGCACATTT 941
Qy 585 TGTAGCCTGTGGAAACGGGGCTTTTTCATCAATTTGCACTTACCTACATTTGAAATTTGACATC 644
Db 942 TGTATGTCTGTGGAGTGGCGCTTTCAGTCTCTGTACTTACTTGAACAGAGGAGGA 1001
Qy 645 ATCTTGAGGACAATATTTTAAAGCTGGAGAACTCACATTTTGAACACGGCCGTGGGAAGA 704
Db 1002 GATCAGAGGACCAAGTTTTCATGATTG---ACTCCAAGTGTGAATCTCGAAAAGACGCT 1058
Qy 705 GTCCATATGACCCCTAAGCTGTGACAGCATCCCTTTTAAATAGATGGAAATTTATATCTGT 764
Db 1059 GCTCTTTCAACCCCAACGTAACCGGTGCTGTTATGATCAATGAGAGGCTTTTCTCTG 1118
Qy 765 GAACTGCAGCTGATTTTATGGGGCGAGACTTTTGTATCTTCCGAACCTTTGGGCAACACC 824
Db 1119 GAATGTATATAGATTTTCATGGGACAGATGCTGCTATTTTTCGAAGTTTAAACCAAGAGA 1178
Qy 825 ACCCAATCAGACAGACAGCATGATTCAGGCTGGCTCAATGATCCAAAGTTTCATTAGTG 884
Db 1179 ATGCGGTGAGAACTGATCAACATAATTCCAAATGGCTAAGTGAACCTATGTTTGTAGATG 1238
Qy 885 CCCACCTCATCTCAGAGAGTGACAATCTGAAGATGACAAAAGTATACCTTTTCTTCCGTG 944
Db 1239 CACATGTCATCCAGATGGTACTGATCCAAATGATGCTAGGTGACTTCTTCTTCAAG 1298
Qy 945 AAAATGCAATAGATGGAGAACACTCTGGAAAAGTACTCAGCGTAGAATAGGTTCAGATAT 1004
Db 1299 AAAAATGACTGACAATAACAGGAGACGAAACAGATTCATTCCATGATGCTCGAATAT 1358
Qy 1005 GCAAGATGACTTTGGAGGCGACAGAGTCTGTTGAATTAATGAGACACATCTCTCAAAG 1064
Db 1359 GTCTTAATGACACTGGTGGACTGCGTAGCCCTTGTCAACAAAGTGGACCACTTCTTAAAGG 1418
Qy 1065 CTCGTCTGATTTGCTCAGTGCAGTCCAAATGGCAATGACACTCATTTTGTATGAATGCTC 1124
Db 1419 CGAGCTGGTGTGCTCGGTAACAGATGAAGACGCGCCCAAGAAACACACTTTGATGAATAG 1478
Qy 1125 AGGATGATTCCTAATGAACCTTTTAAAGATCCTAAAAATCCAGTTGATATGAGAGTTTAA 1184
Db 1479 AGGATGTTTCTGCTGGAACCTGATAACCCGAGGACAACTAGTGTATGGCATTTTAA 1538
Qy 1185 CGACTTCAGTAAACATTTTCAAGGATCAGCCGTGTGTATGTATAGCATGATGATGTA 1244
Db 1539 CAACATCAAGCTCAGTTTTCAAAGGATCAGCCGTGTGTGTATCATTTATCTGTATATAC 1598
Qy 1245 GAAGGCTGTTCTTGGTTCATATATCCACAGGGATGGACCAACTATCAATGGGGGCTT 1304
Db 1599 AGACTGTGTTTAAATGGGCTTTTGGCCACAAGAGGGCCCAATCATCAGCTGATTTCT 1658
Qy 1305 ATCAAGGAAGAGTCCCTTATCCACGGCCAGGAACCTTGTCCAGCAAAAACATTT---GGTG 1361
Db 1659 ATCAGGCGAGAAATTCATATCTCGCCCTGGAACTTGTCCAGGAGGAGCAATTACACCCA 1718
Qy 1362 GTTTTGTACTTCAAAAGGACCTTCTGATGATGTTTAAACCTTTTGAAGAAGTATCCAG 1421
Db 1719 ATATGCGAACCACCAAGGAGTTCAGATGATGTTGTCACTTTTATTCGGAACCATCTCTC 1778

Qy	1422	CCATGTTACAATCCAGTGTCTTCTATGAACAATCGCCCAATAGTGTATCAAAACGGATGTAA	1481
Db	1779	TCATGTTACAATTCATCTACCCAAATCACAAAAGGCCCTTTGATTTGTTTCGTTATTTGGCACTG	1838
Qy	1482	ATTATCAATTTACACAAATTGCTGATACACGAGTGGATGCAGAAGATGGACAGTATGATG	1541
Db	1839	ACTACAGTATACAAGATAGCTGTGATCGATGTAACGCTGCTGATGGAGATATCCATG	1898
Qy	1542	TTATGTTTATCGGAACAGATGTTTGGGACCGTCTTTAAAGTAGTTTCAATTTCTCTAAGGAGA	1601
Db	1899	TCCTGTTTCTCGGAACAGATCGGGGTACTGTGCAAAAGTGGTGTGTTCTTCTCTACTAAACA	1958
Qy	1602	CTTGGTATGATTAGAGAGAGTCTGCTGCGAAGAAATGACAGTTTTTTCGGGAACCGACTG	1661
Db	1959	---ACTCTGTCAGTGGCGAGCTCAATCTGAGGAGCTGGGAAGTCTTTAAAGAAATCATGCTC	2015
Qy	1662	CTATTTCAGCAATGGAGCTTTCCACTAAGCAGCAACAACTATATTTTGGTPTCAACGGCTG	1721
Db	2016	CTATAACAACAATGAAATTTTCACTTAAANAAGCAACAGTTGTATGTGAGTTCCAATGAAG	2075
Qy	1722	GGGTTGCCAGCTCCCTTTACACCGTGTGATATTTACGGGAAAGCGTGTGCTGAGTGTT	1781
Db	2076	GGGTTTCCCAAGTATCTCTGCACCGCTGCCACATCTATGTTACAGCGTGTGCTGACTGCT	2135
Qy	1782	GCCTCGCCCGAGACCCCTTACTGTGCTTGGGATGGTTCTGCATGTTCTCGCTATTTTCCCA	1841
Db	2136	GCCTGGCGCGGACCCCTTTTGGCCTGGGATGGCCATTCCTGTTTCAGATTTCTACCCAA	2195
Qy	1842	CTGCAAAAGAGACGCACAAGACGACAAAGATATAAGAAATGGAGACCCCACTGACTGTT	1901
Db	2196	CTGGGAAACGGAGGAGCCGAAGACAAGATGTGACATGTGAAACCCCACTGACTCAATGCA	2255
Qy	1902	CAGACTTTACACCATGATTAATCAACATGGCCACAGCCCTGAAGAGAGATCATCTATGGTG	1961
Db	2256	GAGGATTT-----TAATCTTAAAGCATACAGAAATGCAGCTGAAATTTGTCAGATATGGAG	2309
Qy	1962	TAGAGAATAGTAGCACATTTTGGAAATGCAGTCGGAAGTCGACAGAGAGCGCTGGTCTATT	2021
Db	2310	TAAAAATACACCACTTTTCTGGAGTGTGCCCCCAAGTCTCCGAGGCACTATCAAGT	2369
Qy	2022	GGCAATTTCCAGAGCGGAAATGAAGAGCGAAAGAAAGAGATCAGAGTGGATGATCATATCA	2081
Db	2370	GGCTGTTTACAGAA--AAGACAAGACAGGAGGAAAGAGGTTAAGCTGAATGAACGAATAA	2426
Qy	2082	TCAGGACAGATCAAGGCCCTTCTGCTAGTGTCTTAACAAGAGGATTCAGGCAATTACC	2141
Db	2427	TAGCCACTTCACAGGGACTCCTGATCGCTCTGTTCAGGGTCTCTGACCAAGGACTTTATC	2486
Qy	2142	TCCTGCCATGGTGGAAACATGGGTTCATACAAACTCTTCTTTAAGGTAAACCTCGGAAGTCA	2201
Db	2487	ACTGCAATGCTACAGAAAATAGTTTTCAAGCAGACCATAGCCCAAGATCACTTCAGAGTTTT	2546
Qy	2202	TTGACACAGAGCAATTGGG	2219
Db	2547	TAGATTCAGAAATGGTGG	2564

RESULT 7

```

US-101-357-820-27
; Sequence 27, Application US/10357820
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.;
; APPLICANT: Burgess, Catherine E.;
; APPLICANT: Casman, Stacie J.;
; APPLICANT: Gorman, Linda;
; APPLICANT: Ji, Weizhen;
; APPLICANT: Kekuda, Ramesh;
; APPLICANT: Li, Li;
; APPLICANT: Padigaru, Muralidhara;
; APPLICANT: Patturajan, Meera;
; APPLICANT: Pena, Carol E.A.;
; APPLICANT: Shenoy, Suresh G.;
; APPLICANT: Shinkets, Richard A.;

```

```

; APPLICANT: Stone, David J.;
; TITLE: Taupier, Raymond J.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING
; FILE REFERENCE: 21402-538B
; CURRENT APPLICATION NUMBER: US/10/357,820
; CURRENT FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: 09/679460
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: 09/730617
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: 10/074978
; PRIOR FILING DATE: 2002-02-12
; PRIOR APPLICATION NUMBER: 10/138588
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 60/387002
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/355099
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/375579
; PRIOR FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/393265
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 401825
; PRIOR FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: 60/381666
; PRIOR FILING DATE: 2002-05-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 27
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (85)..(1626)
; US-10-357-820-27

Query Match          15.1%; Score 407.8; DB 6; Length 1671;
Best Local Similarity 57.4%; Pred. No. 8.7e-121;
Matches 815; Conservative 0; Mismatches 592; Indels 12; Gaps 0

QY 288 CAAGGCTGAAATTATCTCTACAAAGAAATGTTGGAATCCAAACAATGTGATCATTCTTCA
DB 164 CAAGAGTTTATTAAACATTTGATGAACTTCGAGAAACCAAGACCTCTGAATACTTCA

QY 348 GTTTGGCCACAGCTCCAGTTATCATACCTTCCTTTTGGATGAGGAACGGAGTAGGCG
DB 224 TTTCCCAACCATCCTCTTAGACTACAGGATTTTATTAATGATGAAGATCAGGACCGGAG

QY 408 ATGTTGGAGCAAGGATCACATATTTTCATTGCACTGGTTAATATCA--AGGATT
DB 284 ATGTGGAGCAAGATCACATCTTTCCCTGAAATATTAACAATATTAAGTCAAGAAG

QY 465 AAAAGATTGTGTGCCAGTAGTCTTTACACCAAGAAGAGATGAATCAAGTGGCGTGGAA
DB 344 TGAGTGTCTTGCCAGCATCTACAATCAAAGTTGAAGATGCAAAATGCGTGGCA

QY 525 ACATCTGGAAGAAATGTGCTAATTTTCATCAAGGTACTTAAAGGCATATTAATCAGACTCT
DB 404 ATCCCAACACACGGCTGTGGGAACCTTTGTCCGTGTAATTCAGACTTTTCAATCGCACAC

QY 585 TGTACGCGCTGTGAACGGGGGCTTTTCATCCAATTTGCACTTACATTGAAATTTGAGAC
DB 464 TGTATGTCTGTGGAGTGGCGCTTTTCAGTCTGTCTGTACTTACTTTGAACAGAGGAC

QY 645 ATCTCAGGACAATATTTTTTAAGCTGGAACATCTCAATTTTGA AAAACGGCGCTGGGAG
DB 524 GATCAGAGGACCAAGTTTTCATGATTTGA---CTCCAAGTGTGATCTGGAAAAGGAC

QY 705 GTCCATATGACCTTAAGCTGCTGACAGCATCCCTTTTAAATAGATGAGAAATTATATCT
DB 581 GCTCTTTCAACCCCAACGTGAACACGGGTCTGTGTATGATCAATGAGGAGCTTTCTCT

```

```
QY 765 GAACTCAGCTGATTTTATGGGCGAGACTTTGCTATCTTCGAACTCTTTGGGCAACC 824
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 641 GAATGTATATAGATTTTCATGGGCGAGAGCTGCTATTTTTCGAGTTTAAACAAGAGA 700
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 825 ACCAATCAGGACAGACGATGATTCAGGTGGCTCAATGATCCAAAGTTTCATTAGTG 884
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 701 ATGGGTGAGAACTGATCAACATAATTCAAATGGCTAAGTGAACCTATGTTTGTAGATG 760
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 885 CCCACCTCATCTCAGAGAGTGACATCTCTGAAGATGACAAAGTATATCTTTCTCCGTTG 944
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 761 CACATGTCATCCAGATGCTACTGATCCAAATGATCTAAGGTGACTTCTTCTTCAAG 820
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 945 AAAATGCAATAGATGAGAACACTCTGGAAGAGCTACTCACGCTAGAATAGTTCAGATAT 1004
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 821 AAAAAGTACTGACATAACAGGAGCAGCAACAGATTCATTCATGATGCTCGAATAT 880
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1005 GCAAGAACTACTTTGAGGGGCAAGAACTCTGGTGAATAAATGGACAACATTCCTCAAG 1064
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 881 GTCCTAATGACACTGGTGGACTGCGTAGCCTTGTCAACAAGTGGACCACTTCTTAAAGG 940
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1065 CTGCTCTGATTTGCTCAGTGGCAGGTCCAAATGGCAATTCACACTCATTTTGTAGTACTGC 1124
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 941 CGAGGCTGGTGTCTCGTAAACAGATGAAGACGGCCCAAGAAACACACTTTGTATGATTA 1000
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1125 AGGATGATTTCTTAATGAATCTTAAAGATCCTAAATCCAGTTGTATATGAGTGTATA 1184
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1001 AGAATGTGTTCTGCTGGAACCTGATAACCGAGGACACACTAGTGTATGCAATTTTA 1060
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1185 CGAATTCAGTAACATTTTCAAGGGATCAGCCGTGTGTATGATATAGCATGATGTGA 1244
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1061 CAACATCAAGCTCAGTTTTCAAGGATCAGCCGTGTGTATCATTTATCTGTATATAC 1120
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1245 GAAGGTGTTCTTGTTCATATGCCCCACAGGATGGACCCAACTATCAATGCGTGCCTT 1304
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1121 AGACTGTGTTTAAATGGCCCTTTTGCCCAACAAAGAGGGCCCAATCATCAGCTGATTCCT 1180
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1305 ATCAAGGAAGAGTCCCTATCCAGGCGCAGGAACCTTGCCCAACAAACATTTT---GGTG 1361
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1181 ATCAGGCGAGAAATTCATATCTCGCCCTGGAACTTGTCCAGAGGAGCATTTACACCA 1240
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1362 GTTTTGAATCTCAAGAGGACCTTCTCTGATGATGTTATATACCTTTGCAAGAAGTCATCCAG 1421
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1241 ATATGGAACCAACCAAGGAGTTCCACAGATGATGTTGTCACTTTTATTCGGAACCATCTC 1300
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1422 CCAATGTAATCAGTGTGTTCTATGAACAATCGCCCAATAGTATGATCAAAACGGATGTAA 1481
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1301 TCATGTACAAATTCATCTACCAATCCAAAGAGGCTTTGATTTGTTTCTGATTTGGCACTG 1360
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1482 ATTATCAATTTACAAAATTTGTGTAGCCGAGTGGATGCAGAGATGGACAGTATGATG 1541
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1361 ACTACAGTATACAAAGATAGTGTGATGAGTGAAGTGAACGCTGCTGATGGAGATACCATG 1420
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1542 TTATGTTTATCGGAACAGATGTTGGGACCGTTCTTAAAGTAGTGTTCATTTCTTAAGGAGA 1601
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1421 TCTGTTTCTCGGAACAGATCGGGTACTGTGCAAAAAGTGTGTTCTTCTACTAACA 1480
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1602 CTTGATGATTTAGNAGAGTTCTGCTGGAAGAAATGACAGTTTTTTCGGGAACCGACTG 1661
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1481 ---ACTCTGCTAGTGGCAGCTCATTTCTGGAGAGCTGGAAGTCTTTAAGATATCATGCTC 1537
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1662 CTATTTTCAGAAATGAGCTTTTCCACTAAGCAGCAAC 1700
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1538 CTATACAAATGAAATTTTATCTTAAAGAGGAGC 1576
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 8

```
PCT-US03-09929-25
; Sequence 25, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
```

```
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraSeqIist version 0.1
; SEQ ID NO 25
; LENGTH: 1438
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1438)
PCT-US03-09929-25
```

Query Match 4.7%; Score 126.6; DB 1; Length 1438;

Best Local Similarity 50.6%; Pred. No. 4e-30;

Matches 489; Conservative 0; Mismatches 439; Indels 39; Gaps 6;

```
QY 689 AACGCCGTGGAGAGTCCATATGACCTAAGCTGCTGACAGCATCCCTTTTATATAGAT 748
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 749 GGAGAAATATATCTCTGGAACCTGCAGCTGATTTTATGGGCGAGAGCTTTGCTATCTTCCGA 808
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 494 GGAATACTATCTAGCCACAGTCACTGACTTCTTGGCCATTGACGCACTCATTTACCGG 553
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 809 ACTCTTGGGCAACCAACCCCAATCAGGACAGAGCATGATTCAGGTGGCTCAATGAT 868
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 554 AGTCTTGGAGAAAGCCCTTACCCTGCGGACCGTCAAGCACGATTCAAAATGGTTGAAAGAA 613
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 869 CCAAGTTTCATTAGTGGCCACCTCATCTCAGAGAGTGACATCTCGAAGATGACAAAGTA 928
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 614 CCATCTTTGTTCAAGCGGTGGA-----TTACGGAGATTATATC 652
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 929 TACTTTTCTTCCGTGAAAATGCAATAGATGGAGAACACTCTGGAAAAAGCTACTCACGCT 988
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 653 TACTTCTTTCAGGAAATAGCAGTGGAGTATAACACCATGSGAAGTAGTTTTCCTCA 712
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 989 AGAATAGTCAAGATATGCAAGAAATGACTTTGGAGG---GCACAGAGTCTGGTGAATAA 1045
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 713 AGAGTGGCTCAGGTTTGTAAAGATGATATGAGGAGATCTCAAGAGTCTCTGGAGAAACAG 772
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1046 TGGACACATTCCTCAAGCTCGTCTGATTTGCTCAGTCCAGGCTCCAAATGGCATTCAC 1105
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 773 TGGACGCTGTTCTTGAAGGCGCGCTTGAATTTGCTCAGTTCTCTGG-----AGACTCTCAT 826
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1106 ACTCATTTTGTAGTGAACCTGCAGGATGATTTCCTAATGAACCTTTAAAGATCTCTAAATAATCCA 1165
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 827 TTTTATTCAACATTCCTCCAGGCACTTACAGATGTGATTCGTATCAACGGGCGTAT--- 883
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1166 GTTGTATAGAGTGTGTTTACGACTTCCAGTAAACATTTTCAAGGGATCAGCGGTGTATG 1225
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 884 GTTGTCTGGCAACGTTTTTCTACACCTTTATAACAGCATCCCTGGGTCTGCAGTCTGTGCC 943
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

1226 TATAGCATGAGTGTGAGAGGGTTCCTTGTGTCATATGCCACAGGAGTGGACCC 1285
1286 AACTATCAATGGTGC---TTATCAAGAGAGTCCCTATCCACGCGCAGGAATTGT 1342
1004 GATTCACCTGGACACAGTTCCTGATGACGAGTTCCTAAGCCGAGGCGAGTTGCTGT 1063
1343 CCCAGCAA---ACATTTGGTGGTTTGTACTACAAAGGACCTTCCTGATGATGTATA 1399
1064 GCTGCTCATCTCTTAGAAGATATGCAACCTCCAATGAGTTCCTGATGATACCCCTG 1123
1400 ACCTTTGCAAGAGTATCCAGCCATGTACAAATCCAGTGTTCCTATGAAACAATGCCCA 1459
1124 AACTTTCAATCAAGACGACCCGCTCATGGATGAGGAGTCCCTCCATCTTCAACAGGCCA 1183
1460 ATAGTGATCAAAACGAGTGAATTAATCAATTTACACAAATTTGCTAGACCGAGTGAT 1519
1184 TGGTTCCTGAGAACAAATGTCAGATACCGCTTACAAAATTTGCGATGACACAGCTGCT 1243
1520 CGAAGATGACAGTATGATGTTATGTTTATCGGAACAGATGTTGGGACCGTTCCTAAA 1579
1244 GGGCCATATCAGAAATCACTGTGTTTCTGGGATCAGAAAGGATCATCTTGAAG 1303
1580 GTAGTTTCAATTCCTAAGAGACTTGGTATGATTTAGAGAGGTTTCTGCTGGAAGAAATG 1639
1304 TTTTGGCCAGAAATAGGAATAGTGGTTTCTTAAATGACAGCCTTTTCTGGAGGATG 1363
1640 ACAGTTT 1646
1364 AGTGT 1370

RESULT 9

PCT-US03-09929-55
; Sequence 55, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-5738-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 55
; LENGTH: 2113
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1998)

PCT-US03-09929-55

Query Match 4.7%; Score 126.6; DB 1; Length 2113;
Best Local Similarity 50.6%; Pred. No. 4.9e-30;
Matches 489; Conservative 0; Mismatches 439; Indels 39; Gaps 6;
QY 689 AACGCCGTGGAGAGTCCATATGACCCCTAAGCTGTGACAGCATCCCTTTTATAGAT 748
Db 493 AGCGGAATGGCCAGATGCCCATATGATGCCAAACATGCAACGTTTGGCACTGTTTGCAGAT 552
QY 749 GGAGAAATATATCTCTGGAACTGCACTGATTTTATGGGCGAGACTTGTCTATCTTCCGA 808
Db 553 GGAATACTATCTACTGACCCACAGTACTGCTTCTTGCATGACGAGTCAATTTACCG 612
QY 809 ACTCTGGGACACCCACCCCAATCAGGACAGACAGTATTCAGGTGGCTCAATGAT 868
Db 613 AGTCTTGAGAAAGCCCTACCCTGCGGACCGTCAAGCAGATTCAAAATGTTGAAGAA 672
QY 869 CAAAGTTTATGATGCCCCCTCATCTCAGAGAGTGAACAATCTCGAAGATGACAAAGTA 928
Db 673 CCATACCTTTGTTCAAGCCGTGA-----TTACGGAGATTATATC 711
QY 929 TACTTTTCTTCCGTGAAATGCAATAGATGGAGAACACTCTGGAAAGCTACTCACGCT 988
Db 712 TACTTTTCTTTCAGGGAATTCAGTGGAGTATACCAATGGGAAAGGTAGTTTTCCTCA 771
QY 989 AGAATAGTCAAGATATGCAAGATGACCTTTGAGG---GCACAGAACTCTGGTGAATAA 1045
Db 772 AGAGTGGCTCAGTTTGTGAAGATGATGGAGATCTCAAGAGTCTCGGAGAAACAG 831
QY 1046 TGGACAACTTCTCAAAGCTCGTCTGATTTGCTCAGTGGCAGGTCCAAATGGGCAATGAC 1105
Db 832 TGGACGTCTTCTGAAGGCGCTGAACTGCTCAGTTCCTG-----AGACTCTCAT 885
QY 1106 ACTCATTTTGTGAACTGCAGGATGATTTCTTAATGAATCTTAAAGATCTTAAATCCA 1165
Db 886 TTTTATTTCAACATTTCTCAGGAGTATACAGATGATTCGTATCAACGGGCGTGTAT--- 942
QY 1166 GTTGATATGAGTGTTTTACGACTTCCAGTAACTATTTCAAGGGGATCAGCGTGTGTATG 1225
Db 943 GTTGTCTGGCAACGTTTCTACACCTTATAACAGCATCCCTGGTCTGCGAGTCTGTGCC 1002
QY 1226 TATAGCATGAGTGTGAGAGGGTGTTCCTTGGTTCATATGCCCAAGGAGTGGACCC 1285
Db 1003 TATGACATGCTTGACATTTGCCAGTGTTTTACTGGGAGATTCAGAGAAACAGAACTCTCT 1062
QY 1286 AACTATCAATGGGTGCC---TTATCAAGGAGAGTCCCTATCCACGCGCAGGAATTGT 1342
Db 1063 GATTCACCTGGACACCAAGTTCCTGATGAACGAGTTCCTAAGCCGAGGCTGCTGT 1122
QY 1343 CCCAGCAA---ACATTTGGTGGTTTGTGATTTTATCGAAAGGACCTTCCTGATGATGTATA 1399
Db 1123 GCTGCTCATCTCTTAGAAGATATGCAACCTCCATGAGTTCCTGATGATACCCCTG 1182
QY 1400 ACCTTTGCAAGAGTCAATCCAGCCATGPAACAATCCAGTGTTCCTATGAAACAATGCCCA 1459
Db 1183 AACTTCATCAAGACGACCCGCTCATGATGAGGAGTGGCCTCCATCTTCAACAGAGCCA 1242
QY 1460 ATAGTGATCAAAACGAGTGAATTAATCAATTTACACAAATTTGCTAGACCGAGTGTAT 1519
Db 1243 TGGTTCCTGAGAACAAATGCTGATACCGCTTACAAAATTTGCAAGTGGACACAGCTGCT 1302
QY 1520 GCAGAGATGACAGTATGATGTTTATCGAAACAGATGTTGGGACCGTTCCTTAAA 1579
Db 1303 GGGCCATATCAGATCAGTGTGTTTCTGGATCAGAGAGGGAATCATCTTGAAG 1362
QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAGAGGTTTCTGCTGGAAGAAATG 1639
Db 1363 TTTTGGCCAGAAATAGGAATAGTGGTTTCTTAAATGACAGCCTTTTCTCGAGGAGATG 1422
QY 1640 ACAGTTT 1646
Db 1423 AGTGT 1429

Query Match	4.6%;	Score 125;	DB 1;	Length 1327;
Best Local Similarity	50.5%;	Pred. No. 1.3e-29;		
Matches	488;	Conservative 0;	Mismatches 440;	Indels 39; Gaps 6;

QY	689	AACGGCCGTGGGAAGAGCTCCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAATAGAT	748
Db	323	AGCGGAATGCCAGATGCCATATGATGCCAAACATGCCAACGTTGTCAGCTGTTTGCGAGAT	382
QY	749	GGAGAATTACTCTCGAACTGCGAGCTGATTTTATGGGGCGAGACTTTGCTATCTTCCGA	808
Db	383	GGAAACTATCTCAGCCACAGTGACTGCTTCCTGCCATTGACGCAGTCATTTACCGG	442
QY	809	ACTCTTGGGCACCACCAACCCAACTCAGACACAGCAGCATGATTCACAGTGGCTCAATGAT	868
Db	443	AGCTTTGGGAAGGCCCTACCTCGGCACCGTCAAGCAGCATTCAAATGGTTGARAGAA	502
QY	869	CCAAAGTTCAATTAGTGGCCCACTCATCTCAGAGAGTGACAATCTGTAAGATGACAAAGTA	928
Db	503	CCATACTTTGTTCAAAGCCGTGGA-----TTACGGAGATTATATC	541
QY	929	TACTTTTCTTCCTCGTGAAATGCATATAGATGGAGNACACTCTGGAAGACTACTCAGCT	988
Db	542	TACTTCTTCTTCAGGGAATATAGCAGTGGAGTATAACACCATTGGAAAGGTAGTTTTCCCA	601
QY	989	AGAATAGGTCCAGATATGCAAGAATGACTTTGGAGG---GCACAGAACTCTGGTGAATAAA	1045
Db	602	AGAGTGGCTCAGGTTGTAGAAATGATATGGAGGATCTCAAGAGTCTCTGGAGNAACAG	661
QY	1046	TGGACAACATTCCTCAAAGTCGTCTGATTTGCTCAGTGCAGGTCCAAATGGCAATTGAC	1105

```

RESULT 11
PCT-US03-09929-23
Sequence 23, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS OF USE
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
Remaining Prior Application data removed - See File wrapper or PALM.

```

NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 23
; LENGTH: 1492
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1492)
PCT-US03-09929-23

Query Match 4.6%; Score 125; DB 1; Length 1492;

Best Local Similarity 50.5%; Pred. No. 1.3e-29;
Matches 488; Conservative 0; Mismatches 440; Indels 39; Gaps 6;

QY 689 AACGCCGCTGGGAAGAGTCCATATACCTTAAGCTGCTGACAGCATCCCTTTTAATAGAT 748
DB |||||
QY 488 AGCGGAATGCCAGATGCCATATATGATGCAACATGCCAAGTTCACATGTTGCAGAT 547
DB |||||
QY 749 GGAGAAATATATCTCTGGAACTCGAGCTGATTTTATGGGCGAGACTTTGCTATCTTCCGA 808
DB |||||
QY 548 GGAAATATATCTCTGACCCACAGTACTGCTTCTTGCCTTGAAGCAGTCAATTTACCGG 607
DB |||||
QY 809 ACTCTTGGSCACCAACCCATCAGGACAGCAGCATGATTCAGGTGGCTCAATGAT 868
DB |||||
QY 608 AGCTTGGGAAGCCCTTACCCTGCGGACCGTCAAGCAGCATTCAAAATGGTTGAAGAA 667
DB |||||
QY 869 CCAAAATTCATTTAGTGCCCACTCATCTCAGAGAGTGACATCTTGAAGATGACAAAGTA 928
DB |||||
QY 668 CCATCTTTTCAAGCCGTGA-----TTACGGAGATTATATC 706
DB |||||
QY 929 TACTTTTCTCCGTGAAATGCAATAGATGAGAGCACTCTGGAAGCTACTCACGCT 988
DB |||||
QY 707 TACTTCTTCTCAGGGAATAGCAGTGGAGTAAACACCATGGGAAGGTAGTTTTCCTCA 766
DB |||||
QY 989 AGAATAGTTCAGATGCAAGATGACTTTGAGG---GCACAGAGTCTGTGATGATAA 1045
DB |||||
QY 767 AGAGTGGCTCAGGTTTGAAGATGATATGGAGGATCTCAAGAGTCTCTGGAGAAACAG 826
DB |||||
QY 1046 TGGCAACATTCCTCAAAGCTCTGCTGATTTGCTCAGTGCCAGGTCCAAATGGCATGAC 1105
DB |||||
QY 827 TGGAGTCTGCTCTGAGCGCGCTGAACTGCTAGTCTCTGCTG-----AGACTCTCAT 880
DB |||||
QY 1106 ACTCAATTTGATGAATGAGGATGATTTCTTAATGAATTTAAAGATCTTAAATATCCA 1165
DB |||||
QY 881 TTTTATTTCAACATTTCTCAGGAGTACAGATGATGATTCATCAACGGGCGTGAT--- 937
DB |||||
QY 1166 GTTGATATGAGTGTTTTACGACTTCCAGTACATTTTCAAGGATCAGCGGTGTGATG 1225
DB |||||
QY 938 GTTGCTCTGGCAACGTTTCTACACTTATACAGCATCCCTGGGTCTGCACTCTGTGCC 997
DB |||||
QY 1226 TATAGCATGATGATGAGAGGGTGTTCCTTGGTTCATATGCCCACAGGATGGACCC 1285
DB |||||
QY 998 TATGATGCTTGAATTCCTGAGTGTTTTACTGGAGATTCAGGAACAGAGTCTCT 1057
DB |||||
QY 1286 AACTATCAATGGTGCC---TTATCAAGGAAGAGTCCCTATCCAGGCCAGGAACCTGT 1342
DB |||||
QY 1058 GATTCACCTGGACACCACTTCTGATGAACGAGTCTTAAAGCCAGCCAGGTTGCTGT 1117
DB |||||
QY 1343 CCCAGCAA--ACATTTGGTGGTTTGAATCTTCAAGAGGACCTTCTGATGATGATATA 1399
DB |||||
QY 1118 GCTGCTCATCTCTCTTGAAGAGATGCAACCTCAATGATGATGATGATGATGATGAT 1177
DB |||||
QY 1400 ACCTTTGCAAGAAGTTCATCCAGCATGATCAATTCAGTGTTCCTTATGAACATCGCCA 1459
DB |||||
QY 1178 AACTTCATCAAGACGACCGCTCATGATGAGGAGTGGCTCCCTCATCTTCAACAGGCA 1237
DB |||||
QY 1460 ATAGTATCAAAACGGATGTAATTTATCAATTTACAAATTTGCTAGACCGAGTGGAT 1519
DB |||||
QY 1238 TGGTCTCTGAGAACAAATGTCAGATACCGCTTTACCAAATTTGAGTGGACACACTGCT 1297
DB |||||
QY 1520 GCAGAGATGGACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1579
DB |||||

DB 1298 GGGCCATATCAGATCACAACCTGCTGCTGCTTTCTGGATCAGAGAGGGAATCATCTTGAAG 1357
QY 1580 GTAGTTTCAATTCCTAAGCAGACTTGGTATGATTTAGAGAGGTTCTCTGGAAGAAATG 1639
DB |||||
QY 1358 TTTTGGCCAGATAGGAATAGTGTCTTTCTAATGACAGCCTTTTCTGGAGGAGATG 1417
DB |||||
QY 1640 ACAGTTT 1646
DB |||||
DB 1418 AGTGTTT 1424

RESULT 12

PCT-US03-09929-33

; Sequence 33, Application PC/TUS0309929
; GENERAL INFORMATION:

; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929

; CURRENT FILING DATE: 2003-04-01

; PRIOR APPLICATION NUMBER: 60/368,996

; PRIOR FILING DATE: 2002-04-01

; PRIOR APPLICATION NUMBER: 60/369,980

; PRIOR FILING DATE: 2002-04-04

; PRIOR APPLICATION NUMBER: 60/370,381

; PRIOR FILING DATE: 2002-04-05

; PRIOR APPLICATION NUMBER: 60/370,969

; PRIOR FILING DATE: 2002-04-08

; PRIOR APPLICATION NUMBER: 60/371,002

; PRIOR FILING DATE: 2002-04-09

; PRIOR APPLICATION NUMBER: 60/372,002

; PRIOR FILING DATE: 2002-04-12

; PRIOR APPLICATION NUMBER: 60/384,297

; PRIOR FILING DATE: 2002-05-30

; PRIOR APPLICATION NUMBER: 60/386,816

; PRIOR FILING DATE: 2002-06-07

; PRIOR APPLICATION NUMBER: 60/389,123

; PRIOR FILING DATE: 2002-06-13

; PRIOR APPLICATION NUMBER: 60/402,207

; PRIOR FILING DATE: 2002-08-09

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 179

; SOFTWARE: CuraseqList version 0.1

; SEQ ID NO 33

; LENGTH: 1878

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (1)..(1878)

PCT-US03-09929-33

Query Match 4.6%; Score 125; DB 1; Length 1878;

Best Local Similarity 50.5%; Pred. No. 1.5e-29;

Matches 488; Conservative 0; Mismatches 440; Indels 39; Gaps 6;

QY 689 AACGCCGCTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGCATCCCTTTTAATAGAT 748
DB |||||
QY 424 AGCGGAATGCCAGATGCCATATGATGCCAATGCAACATGCCAAGTTCACATGTTGCAGAT 483
DB |||||
QY 749 GGAGAAATATATCTCTGGAACTGCACTGATTTTATGGGCGAGACTTTGCTATCTTCCGA 808
DB |||||
QY 484 GGAAATATATCTCTGAGCAGTACTGCTTCTTGGCAATGACCGAGTCAATTTACCGG 543
DB |||||
QY 809 ACTCTTGGSCACCAACCCATCAGGACAGCAGCATGATTCAGGTGGCTCAATGAT 868
DB |||||
QY 544 AGTCTTGGGAAGCCCTTACCCTGCGGACCGTCAAGCAGCATTCAAAATGGTTGAAGAA 603
DB |||||
QY 869 CCAAAATTCATTTAGTGCCCACTCATCTCAGAGAGTGACATCTTGAAGATGACAAAGTA 928
DB |||||
QY 604 CCATCTTTGTTCAAGCCGTGA-----TTACGGAGATTATATC 642
DB |||||
QY 929 TACTTTTCTTCCGTGAAATGCAATAGATGAGAGAACACTCTGGAAAGCTACTCACGCT 988
DB |||||

Db 643 TACTTCTCTCAGGAAATAGCAGTGGAGTATACACCATGGAAAGGTAGTTTCCCA 702
Qy 989 AGAATAGTCAGATATGCAAGATGACTTTGGAGG---GCACAGAGTCTGGTGAATAA 1045
Db 703 AGAGTGGCTCAGGTTTGAAGATGATATGGGAGGATCTCAAGAGTCTCGGAAACAG 762
Qy 1046 TGGACAAATCTCTCAAGCTGCTGATTTGCTCAGTGGCCAGGTCCTCAATGCAATG 1105
Db 763 TGACGCTGCTCTGAAGCGGCTTGAACGCTGCTCAGTTCTCGG-----AGACTCTAT 816
Qy 1106 ACTCATTTTGTATGAACTGAGGATGATTCCTTAATGAATTTAAAGATCCTAAATCCA 1165
Db 817 TTTTATTTCAACATCTCCAGCAGTTACAGATGATTCGATCAACCGGCGTAT--- 873
Qy 1166 GTTGATATGAGAGTGTATGAGTCTCCAGTACATTTTCAAGGATCAGCGGTGTATG 1225
Db 874 GTTGCTGCTGCAACGTTTTCTACACCTTATAACAGCATCCCTGGGTCTGCACTCTG 933
Qy 1226 TATAGCATGATGATGAGAGGTTTCTTGTGCTATATGCCACAGGATGGACCC 1285
Db 934 TATGACATGCTTGACATGCCAGTGTCTTACTGGAGATTCAGGAACAGAGTCTCT 993
Qy 1286 AACTATCAATGGGTGCC---TTATCAAGGAAGAGTCCCTATCCAGGCGGCAACTGT 1342
Db 994 GATTCACCTGGACACCACTCTCTGATGAACGAGTTCTTAAGCCAGGCGAGTTGCT 1053
Qy 1343 CCAGCAAA---ACATTTGGTGGTTTGTACTCTCAAGAGGACTTCTGATGATGATA 1399
Db 1054 GCTGGCTCATCTCTCTTAGAAGATATGCAACCTCCAATGATGCTCCCTGATGATACC 1113
Qy 1400 ACCTTTGCAAGAGTATCCAGCCATGATCAATCCAGTCTTCTTATGAACATCGCCA 1459
Db 1114 AACTTCATCAAGACGCCCGCTCATGGATGAGGAGTCCCTCACTTTCAACAGGCCA 1173
Qy 1460 ATAGTATCAAAACGAGTAAATATCAATTTACAAATTTGCTAGACCGAGTGGAT 1519
Db 1174 TGGTCTCTGAGAACATGTCAGATACCGCTTACCAAAATGTCAGTGGACACAGTCT 1233
Qy 1520 GCAGAGATGGACAGTATGATGTTTATCGGACAGATGTTGGGACCGTCTTAA 1579
Db 1234 GGGCCATATCAGAAATCACACTGTGGTTTTCTGGATCAGAGAGGGAATCATCTTGA 1293
Qy 1580 GTAGTTTCAATCTCAAGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAAGATG 1639
Db 1294 TTTTGGCCAGATAGGAATAGTGGTTTCTAAATGACAGCCTTTCTGGAGGAGATG 1353
Qy 1640 ACAGTTT 1646
Db 1354 AGTGTTT 1360

RESULT 13
PCT-US03-09929-15
; Sequence 15, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: CibaGen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12

; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 15
; LENGTH: 1921
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1920)
PCT-US03-09929-15

Query Match 4.6%; Score 125; DB 1; Length 1921;
Best Local Similarity 50.5%; Pred.No. 1.5e-29;
Matches 488; Conservative 0; Mismatches 440; Indels 39; Gaps 6;
Qy 689 AACGCCCTGGGAAGAGTCCATATCACCTTAAGCTGCTGACAGCATCCCTTTTAATAGAT 748
Db 406 AGCGAATGGCCAGATGCCATATGATGCCAAACATGCCAACGTTGCACTGTTGCAGAT 465
Qy 749 GGAGAAATPATCTCTGGAACTGCACTGATTTTATGGGGCGAGACTTTGCTATCTTCGA 808
Db 466 GGAAACATATCTCAGCCACAGTACTGACTTCTTGCCATTGACGAGTCAATTTACCGG 525
Qy 809 ACTCTGGGACACACCACCAATCAGGACAGAGCATGATTCAGTGGCTGCTCAATGAT 868
Db 526 AGTCTGGAGAAAGCCCTTACCCCTGGGACCTCAAGCACAGTTCAAAATGGTTGAAGAA 585
Qy 869 CCAAGTTCATTAGTGGCCACCTCATCTCAGAGAGTGAACAATCTCTGAAGATGACAAAGTA 928
Db 586 CCATCTTTTCAAGCGGTGA-----TTACGGAGATTATATC 624
Qy 929 TACTTTTTCTTCGTTGAAAAATGCAATAGATGGAGAACACTCTGGAAGAGTACTCACTCG 988
Db 625 TACTTCTCTTCAGGAAATAGCAGTGGAGTATAACACCATGGGAAAGGTAGTTTTCCTCA 684
Qy 989 AGATAGTCTAGATATGCAAGAAATCACTTTGGAGS---GCACAGAGTCTGCTGAATAA 1045
Db 685 AGATGGGCTCAGGTTTGAAGAAATGATGAGGAGATCTCAAGAGTCTCTGGAGAACAG 744
Qy 1046 TGGACAAATTCCTCAAGCTCGTCTGATTTGCTCAGTGGCCAGGTCCAAATGGCAATTGAC 1105
Db 745 TGGAGTCTGTTCTGAGGCGGCTTGAACCTGCTCAGTTCTCTGG-----AGACTCTCAT 798
Qy 1106 ACTCATTTGTATGAATGCAAGATGATTTCTTAATGAATTTAAAGATCTCTAAATCA 1165
Db 799 TTTTATTTCAACATCTCCAGGAGTTACAGATGATTCGTATCAACGGGCGTAT--- 855
Qy 1166 GTTGATATGAGGTTTACACACTTCCAGTAAACATTTTCAAGGATCAGCGGTGTATG 1225
Db 856 GTTGCTGGCAACGTTTCTTACACCTTATAACAGCATCCCTGGGTCTGCACTCTGTGCC 915
Qy 1226 TATAGCATGATGATGAGAAAGGTTTCTCTGCTCATATGCCCACAGGATGGACCC 1285
Db 916 TATGACATGCTTGACATTTGCCAGTGTTTTACTGGAGATTCAAGGAACAGAGTCTCT 975
Qy 1286 AACTATCAATGGGTGCC---TTATCAAGGAAGAGTCCCTATCCAGGCGCAGGAACCTGT 1342
Db 976 GATTCACCTGGACACCACTTCTCTGATGAACGAGTTCCTAAGCCCGCAGGTTGCTGT 1035
Qy 1343 CCCAGCAAA---ACATTTGGTGGTTTGAACCTCTCAAGAGGACCTTCTCTGATGATGATA 1399
Db 1036 GCTGGCTCATCTCTCTTAGAAGATATGCAACCTCCTCAATGATGTTCCCTGATGATACCT 1095
Qy 1400 ACCTTTGCAAGAGTTCATCCAGCCATGTAACAATCCAGTGTTCCTTATGAACAAATCGCCA 1459

Db 1096 AACTTCATCAAGACGCCCGCTCATGATAGGAGTGCCTCATCTTCAACAGGCCA 1155
Qy 1460 ATAGTGCATCAAAACGGATGTAATATCAATTTACACAAATTTGCTAGACCGAGTGGAT 1519
Db 1156 TGGTTCTTCAGAACAAATGGTCAATACCGCTTACCAAAATTTGAGTGGACACAGCTGCT 1215
Qy 1520 GCAGAGATGGACATGATGTTATGTTATCGGAACAGATGTTGGACCGTCTTTAAA 1579
Db 1216 GGGCCATATCAGAAATCAGACTGCTGTTTCTGGGATCAGAGAAGGGAATCATCTTGAAG 1275
Qy 1580 GTAGTTTCAATTCCTAAGAGACTGGTATGATTTTAGAGAGTTCTGCTGGAAGAAATG 1639
Db 1276 TTTTGGCCAGAAATAGGAATAGTGGTTTCTTAAATGACAGCCTTTTCTCGAGGAGATG 1335
Qy 1640 ACAGTTT 1646
Db 1336 AGTGTTT 1342

RESULT 14
PCT-US03-09929-49
; Sequence 49, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 49
; LENGTH: 1948
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)...(1948)
PCT-US03-09929-49

Query Match 4.6%; Score 125; DB 1; Length 1948;
Best Local Similarity 50.5%; Pred. No. 1.5e-29;
Matches 488; Conservative 0; Mismatches 440; Indels 39; Gaps 6;
Qy 689 AACGCCGTGGGAAGATCCATATGACCCATAGCTGCTGACAGCATCCCTTTTAAATAGAT 748
Db 434 AGCGGAATGGCCAGATGCCATATGATGCCAAACATGCGCACTGTTTTCAGAT 493
Qy 749 GCAGAAATATATCTCTGGAATCAGCTGATTTTATGGGCGAGACTTTGCTATCTTCGA 808
Db 494 GGAAACTATATCTAGCCACAGTACTGCTTCTTGCATTTGACGAGTCAATTTACCGG 553

Qy 809 ACTCTTGGGACACCAACCCCAATCAGACAGAGAGATGATTCAGGTGCTCAATGAT 868
Db 554 AGTCTTGGAGAAAGCCCTTACCCTCGGACCCCTCAAGCAGCATCAAAATGGTTGAAAGAA 613
Qy 869 CCAAGTTTCATTAGTGGCCACCTCATCTCAGAGAGTGACATCTCTGAAGATGACAAAGTA 928
Db 614 CCATACCTTTGTTCAAGCCGTGGA-----TTACGGAGATTATATC 652
Qy 929 TACTTTTTTCTTCGCGTGAATGCAATAGATGGAGAACACTCTCTGGAAGAGCTACTCACGCT 988
Db 653 TACTTCTTTCAGGGAATAGCAGTGGAGTATAACACCATGSGAAGTAGTTTTCCTCA 712
Qy 989 AGAATAGTCTAGATATGCAAGAATGACTTTGGAGS---GCACAGAGTCTCGTGAATAAA 1045
Db 713 AGAGTGGCTCAGGTTTGTGAAGATGATATGGAGGATCTCAAGAGTCTCTGGAGAACAG 772
Qy 1046 TGGACAACTTCTCAAGCTCGTCTGATTTGGCTCAGTCCAGGTCCTCAAAATGGGATTCAC 1105
Db 773 TGGACGCTGTTCTTGAAGGCGCTTGAATGCTCAGTTCTCTGG-----AGACTCTCAT 826
Qy 1106 ACTCATTTTGTGAATGCTCAGGATGATTTCTTAATGAACTTTTAAAGATCTCTAAATAATCCA 1165
Db 827 TTTTATTTCAACATTTCTCCAGGAGTTACAGATGATTCGTATCAACGGCGCTGAT--- 883
Qy 1166 GTTGATATAGAGTGTTTACGACTTTCAGATTAACATTTTCAAGGGATCAGCGTGTGTATG 1225
Db 884 GTTGTCTCGGCAAGTGTCTTACACCTTATAACAGCATCCCTGGGTCTGCACTCTGTGCC 943
Qy 1226 TATAGCATGATGATGAGNAGGTTTCTTGGTCCATATGCCACAGGATGGACCC 1285
Db 944 TATGACATGCTTGACATTTGCCAGTGTTTTACTGGGAGATTCAGGAAACAGAAAGTCTCCT 1003
Qy 1286 AACTATCAATGGGTGCC---TTATCAAGGAAGAGTCCCCTATCCACGGCCAGGAACTTGT 1342
Db 1004 GATTCACCTGGACACCAAGTTCCTGTATGAACAGAGTCTTAAGCCAGGAGTGTCTGT 1063
Qy 1343 CCCAGCAAAA---ACATTTGGTGGTTTGACTCTTCAAGAGACCTTCTCTGATGATGTATA 1399
Db 1064 GCTGCTCATCTCTCTTAGAAGATATGCAACCTCCAATGAGTTCCTCTGATGATACCTGT 1123
Qy 1400 ACCTTTGCAGAAAGTCTCAGCCATGTACATCCAGTGTTCCTATGAACAATCGCCCA 1459
Db 1124 AACTTTCATCAAGCGCACCCGCTCATGATGAGGAGTGCCCTCATCTTCAACAGGCCA 1183
Qy 1460 ATAGTATCAAAACGGATGTAATTTATCAATTTTACACAAATTTGCTAGACCGAGTGGAT 1519
Db 1184 TGGTTCCTGAGAACAAATGGTCAAGTACCGCTTACCAAAATTTGAGTGGACACAGTCT 1243
Qy 1520 GCAGAAATGGACATGATGTTATGTTATCGGAACAGATGTTGGACCGTCTTTAAA 1579
Db 1244 GGGCCATATCAGATCAGACTGTGTTTCTGGATCAGAGAAGGGAATCATCTTGAAG 1303
Qy 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAGAGTTCCTCTGGAAGAAATG 1639
Db 1304 TTTTGGCCAGATAGGAATAGTGGTTTCTTAAATGACAGCCTTTTCTCGAGGAGATG 1363
Qy 1640 ACAGTTT 1646
Db 1364 AGTGTTT 1370

RESULT 15
PCT-US03-09929-51
; Sequence 51, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM.nucleic - nucleic search, using sw model

Run on: July 31, 2003, 12:51:05 ; Search time 181 Seconds
(without alignments)
6606.115 Million cell updates/sec

Title: US-09-774-490-1
Perfect score: 2709
Sequence: 1 aatcctttatctatcgatg.....aggcttttttcttaataacc 2709

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/5A COMB.seq:*
2: /cgn2_6/prodata/1/ina/5B COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
5: /cgn2_6/prodata/1/ina/PTUS COMB.seq:*
6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2508.4	92.6	2601	1	US-08-121-713D-53
2	2508.4	92.6	2601	1	US-08-835-268-53
3	2508.4	92.6	2601	2	US-09-060-692-53
4	2508.4	92.6	2601	3	US-08-833-391-53
5	2508.4	92.6	2601	4	US-09-060-610-53
6	2508.4	92.6	2601	5	PCT-US94-10151A-53
7	1415.2	52.2	1481	1	US-08-136-922-1
8	596.4	22.0	2898	4	US-09-308-1798-2
9	95.2	3.5	3692	4	US-09-077-940A-1
10	84	3.1	2433	4	US-09-300-958A-24
11	83.2	3.1	4157	4	US-08-556-422A-1
12	69.8	2.6	2854	1	US-08-121-713D-57
13	69.8	2.6	2854	1	US-08-835-268-57
14	69.8	2.6	2854	2	US-09-060-692-57
15	69.8	2.6	2854	3	US-08-833-391-57
16	69.8	2.6	2854	4	US-09-060-610-57
17	69.8	2.6	2854	5	PCT-US94-10151A-57
18	68.4	2.5	2790	4	US-09-254-594-5
19	68.4	2.5	3432	4	US-09-254-594-4
20	60.6	2.2	3524	4	US-09-077-940A-3
21	58.6	2.2	7218	1	US-08-232-463-14
22	58.4	2.2	3560	1	US-08-121-713D-59
23	58.4	2.2	3560	1	US-08-835-268-59
24	58.4	2.2	3560	2	US-09-060-692-59
25	58.4	2.2	3560	3	US-08-833-391-59
26	58.4	2.2	3560	4	US-09-060-610-59
27	58.4	2.2	3560	5	PCT-US94-10151A-59

28	54.8	2.0	2504	1	US-08-121-713D-63	Sequence 63, Appl
29	54.8	2.0	2504	1	US-08-835-268-63	Sequence 63, Appl
30	54.8	2.0	2504	2	US-09-060-692-63	Sequence 63, Appl
31	54.8	2.0	2504	3	US-08-833-391-63	Sequence 63, Appl
32	54.8	2.0	2504	4	US-09-060-610-63	Sequence 63, Appl
33	54.8	2.0	2504	5	PCT-US94-10151A-63	Sequence 63, Appl
34	54.6	2.0	2787	4	US-09-254-594-2	Sequence 2, Appl
35	54.6	2.0	3195	4	US-09-254-594-1	Sequence 1, Appl
36	48.6	1.8	2670	1	US-08-121-713D-61	Sequence 61, Appl
37	48.6	1.8	2670	1	US-08-835-268-61	Sequence 61, Appl
38	48.6	1.8	2670	2	US-09-060-692-61	Sequence 61, Appl
39	48.6	1.8	2670	3	US-08-833-391-61	Sequence 61, Appl
40	48.6	1.8	2670	4	US-09-060-610-61	Sequence 61, Appl
41	48.6	1.8	2670	5	PCT-US94-10151A-61	Sequence 61, Appl
42	45.4	1.7	2010	3	US-09-240-410-1	Sequence 1, Appl
43	45.4	1.7	2498	3	US-09-041-236-1	Sequence 1, Appl
44	45.4	1.7	2498	4	US-09-771-467C-1	Sequence 1, Appl
45	38.2	1.4	1818	3	US-09-041-236-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-08-121-713D-53
; Sequence 53, Application US/08121713D
; Patent No. 5639856
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713D
; FILING DATE: 13-SEP-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Oeman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2601 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..2331
US-08-121-713D-53
Query Match 92.6%; Score 2508.4; DB 1; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 2509; Conservative		0; Mismatches	1; Indels	0; Gaps	0;
QY	193	CTGCAGCATGGGCTGCTTAAC	TAGGATTCCTGCTCTTTCTGGGAGTATTACTTACAGC	252	
Db	9	CTGCAGCATGGGCTGCTTAAC	TAGGATTCCTGCTCTTTCTGGGAGTATTACTTACAGC	68	
QY	253	AAGAGCAAACTATCAGAAATGGGAAGAA	CAATGCGCAAGCTTGAAATTTATCTTACAAAGA	312	
Db	69	AAGAGCAAACTATCAGAAATGGGAAGAA	CAATGCGCAAGCTTGAAATTTATCTTACAAAGA	128	
QY	313	AATGTTGGAATCAACAATGTGATCACTTT	CAATGGCTTGGCCAAAGCTTCCAGTTATCA	372	
Db	129	AATGTTGGAATCAACAATGTGATCACTTT	CAATGGCTTGGCCAAAGCTTCCAGTTATCA	188	
QY	373	TACCTTCCTTTGGATGAGGAACGGAGTAG	CGCTGATGTTGGAGCAAGGATCACATATT	432	
Db	189	TACCTTCCTTTGGATGAGGAACGGAGTAG	CGCTGATGTTGGAGCAAGGATCACATATT	248	
QY	433	TTCAATTCGACCTGGTTAATATCAAGGAT	TTTCAAAAGATTTGTGTGGCCAGTATCTTACAC	492	
Db	249	TTCAATTCGACCTGGTTAATATCAAGGAT	TTTCAAAAGATTTGTGTGGCCAGTATCTTACAC	308	
QY	493	CAGAAGAGATGAATCAAGTGGGCTCGAA	AGACATCTCTGAAAGATGTCTAAATTTCA	552	
Db	309	CAGAAGAGATGAATCAAGTGGGCTCGAA	AGACATCTCTGAAAGATGTCTAAATTTCA	368	
QY	553	CAAGGTACTTTAGGCAATATCAAGACTCA	CTTGTAGCGCTGTGGACGGGGCTTTTCA	612	
Db	369	CAAGGTACTTTAGGCAATATCAAGACTCA	CTTGTAGCGCTGTGGACGGGGCTTTTCA	428	
QY	613	TCCAAATTCGACCTACATTTGAAATGGA	TCATCATCTGAGGACAAATATTTTAAAGCTGA	672	
Db	429	TCCAAATTCGACCTACATTTGAAATGGA	TCATCATCTGAGGACAAATATTTTAAAGCTGA	488	
QY	673	GAACTCACATTTTGAACCGGCGTGGGA	AGAGTCCATATGACCCCTAAGCTGCTGACAGC	732	
Db	489	GAACTCACATTTTGAACCGGCGTGGGA	AGAGTCCATATGACCCCTAAGCTGCTGACAGC	548	
QY	733	ATCCCTTTTAAATAGATGGAATTTACTCT	GGAAGTCTGGAAGTCTGATTTTATGGGGGAGA	792	
Db	549	ATCCCTTTTAAATAGATGGAATTTACTCT	GGAAGTCTGGAAGTCTGATTTTATGGGGGAGA	608	
QY	793	CTTTGTCTATCTTCGGAATCTTTGGGCA	CCACCAATCAGACAGACAGCATGATTC	852	
Db	609	CTTTGTCTATCTTCGGAATCTTTGGGCA	CCACCAATCAGACAGACAGCATGATTC	668	
QY	853	CAGGTGGCTCAATGATCCAAAGTTTCA	TAGTGCCCACTCATCTCAGAGAGTGACAATCC	912	
Db	669	CAGGTGGCTCAATGATCCAAAGTTTCA	TAGTGCCCACTCATCTCAGAGAGTGACAATCC	728	
QY	913	TGAAGATGACAAAGTATACCTTTTCTTC	CGTGAATGCAATAGATGGAGAACACTCTGG	972	
Db	729	TGAAGATGACAAAGTATACCTTTTCTTC	CGTGAATGCAATAGATGGAGAACACTCTGG	788	
QY	973	AAAAGCTACTCAGCTTAGAATAGTCA	GATATGCAAGATGACTTTGGAGGCAACAGAG	1032	
Db	789	AAAAGCTACTCAGCTTAGAATAGTCA	GATATGCAAGATGACTTTGGAGGCAACAGAG	848	
QY	1033	TCGTGGAATAAATGACAACTTCCTCA	AAAGCTCGTGTGATTTGCTCAGTGCCAGGTCC	1092	
Db	849	TCGTGGAATAAATGACAACTTCCTCA	AAAGCTCGTGTGATTTGCTCAGTGCCAGGTCC	908	
QY	1093	AAATGGCATTTGATGAACTGAGGATG	ATTTTCCCTAAATGAACTTTTAAAGA	1152	
Db	909	AAATGGCATTTGATGAACTGAGGATG	ATTTTCCCTAAATGAACTTTTAAAGA	968	
QY	1153	TCTTAAATTCAGTTGTATAGGAGTGT	TATCCAGTTCAGTAACTTTTCAAGGGATC	1212	
Db	969	TCTTAAATTCAGTTGTATAGGAGTGT	TATCCAGTTCAGTAACTTTTCAAGGGATC	1028	
QY	1213	AGCCGTGTGTATGATAGCATGATGTA	GAAGGGTGTCTCTTGGTCCATATGCCCA	1272	
Db	1029	AGCCGTGTGTATGATAGCATGATGTA	GAAGGGTGTCTCTTGGTCCATATGCCCA	1088	

QY	1273	CAGGATGAGCCCAACTATCAATGGTGC	CTTTATCAAGGAAGAGTCCCTATCCAGGCC	1332	
Db	1089	CAGGATGAGCCCAACTATCAATGGTGC	CTTTATCAAGGAAGAGTCCCTATCCAGGCC	1148	
QY	1333	AGGAACCTTTGTCCAGCAAAACATTT	GGTGTGACTCTACAAAGGACCTTCTGATGA	1392	
Db	1149	AGGAACCTTTGTCCAGCAAAACATTT	GGTGTGACTCTACAAAGGACCTTCTGATGA	1208	
QY	1393	TGTTATAACCTTTGCAAGAAGTCATCC	AGGCCATGTACAAATCAGTGTTCCTATGA	1452	
Db	1209	TGTTATAACCTTTGCAAGAAGTCATCC	AGGCCATGTACAAATCAGTGTTCCTATGA	1268	
QY	1453	TCGCCCAATAGTGTACAAACGAGTAA	ATTTATCAATTTACAAATTTGCTGACCCG	1512	
Db	1269	TCGCCCAATAGTGTACAAACGAGTAA	ATTTATCAATTTACAAATTTGCTGACCCG	1328	
QY	1513	AGTGATGACAGAAGATGACAGTATG	ATGTTATCGGAACAGATTTGGGACCGT	1572	
Db	1329	AGTGATGACAGAAGATGACAGTATG	ATGTTATCGGAACAGATTTGGGACCGT	1388	
QY	1573	TCCTAAAGTAGTTTCAATTCCTAAGGA	CACTTGGTATGATTTAGAAAGGTTCTGCTGA	1632	
Db	1389	TCCTAAAGTAGTTTCAATTCCTAAGGA	CACTTGGTATGATTTAGAAAGGTTCTGCTGA	1448	
QY	1633	AGAAATGACAGTTTTTCGGGAACCG	ACTGCTATTTTTCAGCAATGGAGCTTTCC	1692	
Db	1449	AGAAATGACAGTTTTTCGGGAACCG	ACTGCTATTTTTCAGCAATGGAGCTTTCC	1508	
QY	1693	GCAACAACTATATATTTGCTTCAAC	GGCTTCCAGCTTCCCTTTACACCGGTGA	1752	
Db	1509	GCAACAACTATATATTTGCTTCAAC	GGCTTCCAGCTTCCCTTTACACCGGTGA	1568	
QY	1753	TATTTACGGGAAGGCTGTGAGTGT	TGCTCGCCCGAGACCTTACTGTCTGGA	1812	
Db	1569	TATTTACGGGAAGGCTGTGAGTGT	TGCTCGCCCGAGACCTTACTGTCTGGA	1628	
QY	1813	TGGTTCTCATGTTCTCGCTATTTTCC	CACTTCCAAAGAGAGCAACAGCAAGATAT	1872	
Db	1629	TGGTTCTCATGTTCTCGCTATTTTCC	CACTTCCAAAGAGAGCAACAGCAAGATAT	1688	
QY	1873	AAGAAATGGAGACCCACTGACTCA	CTGTTTACAGACTTACCAATGATATAC	1932	
Db	1689	AAGAAATGGAGACCCACTGACTCA	CTGTTTACAGACTTACCAATGATATAC	1748	
QY	1933	CAGCCCTGAAGAGAGATCATCTAT	GTTGTAGAGATAGTAGCAATTTTGGAA	1992	
Db	1749	CAGCCCTGAAGAGAGATCATCTAT	GTTGTAGAGATAGTAGCAATTTTGGAA	1808	
QY	1993	TCCGAAGTCGAGAGAGCGCTGCT	TATTTGCAATTTCCAGAGGCGAAATGA	2052	
Db	1809	TCCGAAGTCGAGAGAGCGCTGCT	TATTTGCAATTTCCAGAGGCGAAATGA	1868	
QY	2053	AGAAGATCAGAGTGGATGATCAT	TATCATCAGGACAGATCAAGGCTTCTG	2112	
Db	1869	AGAAGATCAGAGTGGATGATCAT	TATCATCAGGACAGATCAAGGCTTCTG	1928	
QY	2113	TCTACACAGAGAGTTTACGCAAT	TACTCTGCCATCGGTGGAAACATGGG	2172	
Db	1929	TCTACACAGAGAGTTTACGCAAT	TACTCTGCCATCGGTGGAAACATGGG	1988	
QY	2173	AACCTCTTAAAGGTTAAACCTTGA	AGTCAATTTGGAAGAACTTCTTCA	2232	
Db	1989	AACCTCTTAAAGGTTAAACCTTGA	AGTCAATTTGGAAGAACTTCTTCA	2048	
QY	2233	TAAAGATGATGAGATGGCTCTA	AGACAAAGAAATGTCCAATAGCATG	2292	
Db	2049	TAAAGATGATGAGATGGCTCTA	AGACAAAGAAATGTCCAATAGCATG	2108	
QY	2293	CCAGAGGTCTGGTACAGAGACTT	CATGACCTCAACCCCAATCTCAAC	2352	
Db	2109	CCAGAGGTCTGGTACAGAGACTT	CATGACCTCAACCCCAATCTCAAC	2168	

Db 969 TCTAAAAATCCAGTTGTATATGAGTGTATTACGACTTCCAGTAAACATTTTCAAGGATC 1028
Qy 1213 AGCCGTGTGTATGTATAGCATGATGTGAGAGGGTGTCTTGTGTCCTATATGSCCA 1272
Db 1029 AGCCGTGTGTATGTATAGCATGATGTGAGAGGGTGTCTTGTGTCCTATATGSCCA 1088
Qy 1273 CAGGATGGAACCCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTATCCAGGCC 1332
Db 1089 CAGGATGGAACCCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTATCCAGGCC 1148
Qy 1333 AGAACTGTGTCAGCAAAACATTTGGTGTGTTGACTCTCAAGAGACCTTCTCTGATGA 1392
Db 1149 AGAACTGTGTCAGCAAAACATTTGGTGTGTTGACTCTCAAGAGACCTTCTCTGATGA 1208
Qy 1393 TGTATAAATCTTGGCAAGAGTCAATCCAGCATGTATCAATCCAGTGTCTTCTATGAACA 1452
Db 1209 TGTATAAATCTTGGCAAGAGTCAATCCAGCATGTATCAATCCAGTGTCTTCTATGAACA 1268
Qy 1453 TGCCCAATAGTATCAAAACCGATGTAAATTTATCAAAATTTGTGCTAGACCG 1512
Db 1269 TGCCCAATAGTATCAAAACCGATGTAAATTTATCAAAATTTGTGCTAGACCG 1328
Qy 1513 AGTGGATGCAAGATGGAAGTATGATGTTATGTTTATCGGAACAGATGTTGGACCGT 1572
Db 1329 AGTGGATGCAAGATGGAAGTATGATGTTATGTTTATCGGAACAGATGTTGGACCGT 1388
Qy 1573 TCTTAAAGTAGTTTCAATTCCTAAGAGAGTGGTATGATTTAGAGAGGTTCTGCTGA 1632
Db 1389 TCTTAAAGTAGTTTCAATTCCTAAGAGAGTGGTATGATTTAGAGAGGTTCTGCTGA 1448
Qy 1633 AGAAATGACAGATTTTTCGGGAACCGACTGTATTTTCAGCAATGGAGCTTTTCCACTAAGCA 1692
Db 1449 AGAAATGACAGATTTTTCGGGAACCGACTGTATTTTCAGCAATGGAGCTTTTCCACTAAGCA 1508
Qy 1693 GCAACAACTATATATTTGGTTCAACGGCTGGGGTGGCCAGCTCCCTTTACCGGTGA 1752
Db 1509 GCAACAACTATATATTTGGTTCAACGGCTGGGGTGGCCAGCTCCCTTTACCGGTGA 1568
Qy 1753 TATTTACGGGAAGCGTGTGCTGAGTGTTCCTCGCCGCGAGACCTTACTGTCTGGA 1812
Db 1569 TATTTACGGGAAGCGTGTGCTGAGTGTTCCTCGCCGCGAGACCTTACTGTCTGGA 1628
Qy 1813 TGCTTCTGATGTTCTCGCTATTTTCCCACTGCAAGAGACGCAAGACGCAAGATAT 1872
Db 1629 TGCTTCTGATGTTCTCGCTATTTTCCCACTGCAAGAGACGCAAGACGCAAGATAT 1688
Qy 1873 AGAAATGAGACCCCACTGACTCACTGTTTCAAGACTTACACCATGATAATCAGCATGCCA 1932
Db 1689 AGAAATGAGACCCCACTGACTCACTGTTTCAAGACTTACACCATGATAATCAGCATGCCA 1748
Qy 1933 CAGCCCTGAAGAGAGATCATCTATGCTGTAGAGATAGTACACATTTTTCGAATGCGAG 1992
Db 1749 CAGCCCTGAAGAGAGATCATCTATGCTGTAGAGATAGTACACATTTTTCGAATGCGAG 1808
Qy 1993 TCCGAAGTCCGAGAGAGCGTGTCTATTTGGCAATTTCCAGAGCGGAAATGAAGAGCGAAA 2052
Db 1809 TCCGAAGTCCGAGAGAGCGTGTCTATTTGGCAATTTCCAGAGCGGAAATGAAGAGCGAAA 1868
Qy 2053 AGAAGAGATCAGAGTGGATGATCATATCATAGGACAGATCAAGGCTTCTCTAGTATG 2112
Db 1869 AGAAGAGATCAGAGTGGATGATCATATCATAGGACAGATCAAGGCTTCTCTAGTATG 1928
Qy 2113 TCTACACAGAGGATTCAGGCAATTTACTCTGCAATGGGTGGAGCATGGGTTCATACA 2172
Db 1929 TCTACACAGAGGATTCAGGCAATTTACTCTGCAATGGGTGGAGCATGGGTTCATACA 1988
Qy 2173 AACTCTTCTTAAGGTAACTTGAAGTCAATGACACAGAGCATTTTGGAAAGAACTTCTTCA 2232
Db 1989 AACTCTTCTTAAGGTAACTTGAAGTCAATGACACAGAGCATTTTGGAAAGAACTTCTTCA 2048
Qy 2233 TAAAGATGATGAGAGTGGCTTAAAGACCAAGAAATGTCCAAATAGCATGACACCTAG 2292
Db 2049 TAAAGATGATGAGAGTGGCTTAAAGACCAAGAAATGTCCAAATAGCATGACACCTAG 2108

Qy 2293 CCAGAAGGTCTGGTACAGAGACTTCATGCGCTCATCAACCAACCCCAATCTCAACACGAT 2352
Db 2109 CCAGAAGGTCTGGTACAGAGACTTCATGCGCTCATCAACCAACCCCAATCTCAACACGAT 2168
Qy 2353 GGATGAGTCTGTGTGAACAAGTTTGAAGAGGACCGAAACCAACGTCGGCAAGGCCAGG 2412
Db 2169 GGATGAGTCTGTGTGAACAAGTTTGAAGAGGACCGAAACCAACGTCGGCAAGGCCAGG 2228
Qy 2413 ACATACCCAGGGAACAGTAAACAATGAGACACTTACAGAGAAATTAAGAAAGGTAGAAA 2472
Db 2229 ACATACCCAGGGAACAGTAAACAATGAGACACTTACAGAGAAATTAAGAAAGGTAGAAA 2288
Qy 2473 CAGGAGGACCCACGAATTTGAGAGGCGCCACGAGGAGTGTCTGAGCTGCAATACCTCTAGA 2532
Db 2289 CAGGAGGACCCACGAATTTGAGAGGCGCCACGAGGAGTGTCTGAGCTGCAATACCTCTAGA 2348
Qy 2533 AACCTCAACAAAGTAGAATACTTGCCTAGACAATACTGGAACCAAAATGCAATATACAT 2592
Db 2349 AACCTCAACAAAGTAGAATACTTGCCTAGACAATACTGGAACCAAAATGCAATATACAT 2408
Qy 2593 GAACCTTTTTCATGCGCATATGTCATGTTTACATGTTGGGAATTCAGCTGAGTTCCA 2652
Db 2409 GAACCTTTTTCATGCGCATATGTCATGTTTACATGTTGGGAATTCAGCTGAGTTCCA 2468
Qy 2653 CCAATTATAAATAAATCAATCCATGAGTAACTTTCTTAATAGGCTTTTTCCTCC 2702
Db 2469 CCAATTATAAATAAATCAATCCATGAGTAACTTTCTTAATAGGCTTTTTCCTCC 2518

RESULT 3

US-09-060-692-53
; Sequence 53, Application US/09060692
; Patent No. 5935865
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,692
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2601 base pairs
; TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: CDS
LOCATION: 16..2331
US-09-060-692-53

Query Match 92.6%; Score 2508.4; DB 2; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	193	CTGCAGCATGGCTGGTAACTAGGATGCTCTGCTCTTTCTCGGAGATTAATCTTACAGC	252
Db	9	CTGCAGCATGGCTGGTAACTAGGATGCTCTCTTTCTCGGAGATTAATCTTACAGC	68
QY	253	AAGAGCAAACTACAGATGGGAAGAAACAATGTGCAAGGCTGAAATATCTCTACAAAGA	312
Db	69	AAGAGCAAACTACAGATGGGAAGAAACAATGTGCAAGGCTGAAATATCTCTACAAAGA	128
QY	313	AATGTTGGAATCAACAATGTGATCACTTTCAATGGCTTGGCCACAGCTCCAGTTATCA	372
Db	129	AATGTTGGAATCAACAATGTGATCACTTTCAATGGCTTGGCCACAGCTCCAGTTATCA	188
QY	373	TACCTTCTTTGGATGAGGAACGAGTAGGCTGTATGTTGAGCAAAAGGATCAATATT	432
Db	189	TACCTTCTTTGGATGAGGAACGAGTAGGCTGTATGTTGAGCAAAAGGATCAATATT	248
QY	433	TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTCGCCAGATCTTACAC	492
Db	249	TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTCGCCAGATCTTACAC	308
QY	493	CAGAAGAGATGAATCAAGTGGGCTGGAAAGACATCTCTGAAAGAAATGTCTAAATTCAT	552
Db	309	CAGAAGAGATGAATCAAGTGGGCTGGAAAGACATCTCTGAAAGAAATGTCTAAATTCAT	368
QY	553	CAAGGTACTTAAGGATATAATCAGATCTACCTGTGAGCAAGGCGGCTTTTCA	612
Db	369	CAAGGTACTTAAGGATATAATCAGATCTACCTGTGAGCAAGGCGGCTTTTCA	428
QY	613	TCCAAATTCGACCTCAATTTGAATTTGGACATCATCTGAGGACATATTTTAAAGCTGA	672
Db	429	TCCAAATTCGACCTCAATTTGAATTTGGACATCATCTGAGGACATATTTTAAAGCTGA	488
QY	673	GAACTCACATTTTGAACCGGCGGTGGGAAGAGTCCATATGACCTCAAGCTGCTACAGC	732
Db	489	GAACTCACATTTTGAACCGGCGGTGGGAAGAGTCCATATGACCTCAAGCTGCTACAGC	548
QY	733	ATCCCTTTTAATAGATGGGAATTAATCTCTGGAACCTGAGCTGATTTTATGGGGCGAGA	792
Db	549	ATCCCTTTTAATAGATGGGAATTAATCTCTGGAACCTGAGCTGATTTTATGGGGCGAGA	608
QY	793	CTTTGCTATCTTCGAACTCTTGGGCGACCATCCACCAATCAGGACAGGACGATGATTC	852
Db	609	CTTTGCTATCTTCGAACTCTTGGGCGACCATCCACCAATCAGGACAGGACGATGATTC	668
QY	853	CAGGTGGCTCAATGATCCAAAGTTCATTTAGTCCCACTCTCTCAGAGAGTGACAATCC	912
Db	669	CAGGTGGCTCAATGATCCAAAGTTCATTTAGTCCCACTCTCTCAGAGAGTGACAATCC	728
QY	913	TGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGGAGAACACTCTGG	972
Db	729	TGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGGAGAACACTCTGG	788
QY	973	AAAAGCTACTCAGCTAGATAGGTCAGATATGCAAGATGATCTTTGGAGGCGACAGAAG	1032
Db	789	AAAAGCTACTCAGCTAGATAGGTCAGATATGCAAGATGATCTTTGGAGGCGACAGAAG	848
QY	1033	TCTGTTGAATAAATGACAACTTCTCAAAGCTCGTCTGATTTGCTCAGTGGCCAGGTCC	1092
Db	849	TCTGTTGAATAAATGACAACTTCTCAAAGCTCGTCTGATTTGCTCAGTGGCCAGGTCC	908
QY	1093	AAATGGCATTTGACATCATTTTGTGAACTGCGAGATGTATTTCTTAATGAATTTAAAGA	1152

Db	909	AAATGGCATTTGACACTCAATTTTGATGAACCTCCAGATGTATTTCTTAATGAATTTAAAGA	968
QY	1153	TCCTAAAAATCCAGTTGTATATGGAGTGTATACGACTTCCAGTAAACATTTTCAAGGGATC	1212
Db	969	TCCTAAAAATCCAGTTGTATATGGAGTGTATACGACTTCCAGTAAACATTTTCAAGGGATC	1028
QY	1213	AGCCGTGTATATATAGCATGATGTGAGAAAGGTGTCTTGTGTCATATGCCCCA	1272
Db	1029	AGCCGTGTATATATAGCATGATGTGAGAAAGGTGTCTTGTGTCATATGCCCCA	1088
QY	1273	CAGGGATGGACCCAACTATCAATGGTGCCTTATCAGGAAGAGTCCCTATCCACGCC	1332
Db	1089	CAGGGATGGACCCAACTATCAATGGTGCCTTATCAGGAAGAGTCCCTATCCACGCC	1148
QY	1333	AGGAATCTGTCTCCAGCAAAACATTTGGTGGTGTGACTCTACAAAGGACCTTCTCTGATGA	1392
Db	1149	AGGAATCTGTCTCCAGCAAAACATTTGGTGGTGTGACTCTACAAAGGACCTTCTCTGATGA	1208
QY	1393	TGTTATAACCTTTCCAAAGATCATCCAGCCATGTACAAATCCAGTGTTCCTATGAACAA	1452
Db	1209	TGTTATAACCTTTCCAAAGATCATCCAGCCATGTACAAATCCAGTGTTCCTATGAACAA	1268
QY	1453	TCGCCCAATAGTGAATCAAAACGGATGTAAATTAATCAATTTACAAATTTGCTAGACCG	1512
Db	1269	TCGCCCAATAGTGAATCAAAACGGATGTAAATTAATCAATTTACAAATTTGCTAGACCG	1328
QY	1513	AGTGGATCGAAGATGAGACATGATGTATGTATTTATCGGAACAGATTTGGGACCGT	1572
Db	1329	AGTGGATCGAAGATGAGACATGATGTATGTATTTATCGGAACAGATTTGGGACCGT	1388
QY	1573	TCTTAAAGTAGTGTCAATTTCTAAAGGACATTTGGTGTATTTAGAAAGGTTCTGCTGGA	1632
Db	1389	TCTTAAAGTAGTGTCAATTTCTAAAGGACATTTGGTGTATTTAGAAAGGTTCTGCTGGA	1448
QY	1633	AGAAATGACAGTTTTTGGGAAACCGATCTGCTATTTTACCAATGAGCTTTTCCACTAAGCA	1692
Db	1449	AGAAATGACAGTTTTTGGGAAACCGATCTGCTATTTTACCAATGAGCTTTTCCACTAAGCA	1508
QY	1693	GCAACAACTATATATTTGGTTCACCGCTGGGGTGGCCAGCTCCCTTTTACACCGGTGGA	1752
Db	1509	GCAACAACTATATATTTGGTTCACCGCTGGGGTGGCCAGCTCCCTTTTACACCGGTGGA	1568
QY	1753	TATTTACGGGAAAGCGTGTCTGAGTGTGCTCCCGGAGACCTTTACTGTCTGGGA	1812
Db	1569	TATTTACGGGAAAGCGTGTCTGAGTGTGCTCCCGGAGACCTTTACTGTCTGGGA	1628
QY	1813	TGGTTCTGCTATGTTCTGCTATTTTCCCATCTGCAAGAGACGCAAGACGACATAT	1872
Db	1629	TGGTTCTGCTATGTTCTGCTATTTTCCCATCTGCAAGAGACGCAAGACGACATAT	1688
QY	1873	AAGAAATGGAGACCCACTGACTCACTGTTCAGACTTACACCATGATATACCATGGCCA	1932
Db	1689	AAGAAATGGAGACCCACTGACTCACTGTTCAGACTTACACCATGATATACCATGGCCA	1748
QY	1933	CAGCCCTGAAGAGAGAAATCATCTATGTTGTAGAGATAGTAGCACATTTTGGAAATGCGAG	1992
Db	1749	CAGCCCTGAAGAGAGAAATCATCTATGTTGTAGAGATAGTAGCACATTTTGGAAATGCGAG	1808
QY	1993	TCCGAAGTGGCAGAGAGCGCTGTCTATTTGGCAATTTCCAGAGCGGAAATGAAGAGCGAAA	2052
Db	1809	TCCGAAGTGGCAGAGAGCGCTGTCTATTTGGCAATTTCCAGAGCGGAAATGAAGAGCGAAA	1868
QY	2053	AGAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGATAG	2112
Db	1869	AGAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGATAG	1928
QY	2113	TCTACACAGAGAGATTTACGCAATTTACCTCTGCGATCGCGTGGAAACATGGTTTCATACA	2172
Db	1929	TCTACACAGAGAGATTTACGCAATTTACCTCTGCGATCGCGTGGAAACATGGTTTCATACA	1988
QY	2173	AACTCTTTTAAGGTAAACCTTGGAAAGTCAATTTGACACAGAGATTTTGGAAAGCTTCTTCA	2232

Db 1989 AACTCTTCTTAAGGTAACTGGAAGTCAATTGACACAGAGCAATTTGGAAAGAACTTTCTTCA 2048
Qy 2233 TAAAGATGATGAGGAGTCTTAAGACCAAGAAATGTCCAATGAGCATGACACCTAG 2292
Db 2049 TAAAGATGATGAGGAGTCTTAAGACCAAGAAATGTCCAATGAGCATGACACCTAG 2108
Qy 2293 CCAGAAAGGTCTGTCAGAGAGACTTCATGAGCTCATCAACCCCAATCTCAACACGAT 2352
Db 2109 CCAGAAAGGTCTGTCAGAGAGACTTCATGAGCTCATCAACCCCAATCTCAACACGAT 2168
Qy 2353 GGATGAGTCTGTCAGAGAGACTTCGAAAGAGGACCAAAACAGCTGGCAAGGCGAGG 2412
Db 2169 GGATGAGTCTGTCAGAGAGACTTCGAAAGAGGACCAAAACAGCTGGCAAGGCGAGG 2228
Qy 2413 ACATACCCAGGAGCAAGTAAACAAATGGAAGCACTTACAGAAATAAAGAAAGTAGAAA 2472
Db 2229 ACATACCCAGGAGCAAGTAAACAAATGGAAGCACTTACAGAAATAAAGAAAGTAGAAA 2288
Qy 2473 CAGGAGGACCCAGAAATTTGAGAGGACCCAGGAGTGTCTGAGCTGCAATTACCTCTAGA 2532
Db 2289 CAGGAGGACCCAGAAATTTGAGAGGACCCAGGAGTGTCTGAGCTGCAATTACCTCTAGA 2348
Qy 2533 AACCTCAACAGTAGAAGTTCGCTAGACATCAATCAATGGAAGAAACAAATGCAATATACAT 2592
Db 2349 AACCTCAACAGTAGAAGTTCGCTAGACATCAATCAATGGAAGAAACAAATGCAATATACAT 2408
Qy 2593 GAACCTTTTTCATGCGCATTTATGAGATGTTTCAATGGTGGGAAATTCAGCTGAGTTCCA 2652
Db 2409 GAACCTTTTTCATGCGCATTTATGAGATGTTTCAATGGTGGGAAATTCAGCTGAGTTCCA 2468
Qy 2653 CCAATTAATAATTAATCCATGAGTAACTTTCTTAATAGGCTTTTTTTTCC 2702
Db 2469 CCAATTAATAATTAATCCATGAGTAACTTTCTTAATAGGCTTTTTTTTCC 2518

RESULT 4

US-08-833-391-53
; Sequence 53, Application US/08833391
; Patent No. 6013781
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,391
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342

TELEX:
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2601 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..2331
; US-08-833-391-53

Query Match 92.6%; Score 2508.4; DB 3; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 193 CTGCAGCATGGCTGGTTAACTAGAGTTGTCTGTTCTTTCTGGGAGATTAATTCTTACAGC 252
Db 9 CTGCAGCATGGCTGGTTAACTAGAGTTGTCTGTTCTTTCTGGGAGATTAATTCTTACAGC 68
Qy 253 AAGACCAACTATCAGATGGGAGAGCAATGTGCCAAGGCTGAAATTTATCTTACAAGA 312
Db 69 AAGACCAACTATCAGATGGGAGAGCAATGTGCCAAGGCTGAAATTTATCTTACAAGA 128
Qy 313 AATGTTGGAATCCAACTATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCA 372
Db 129 AATGTTGGAATCCAACTATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCA 188
Qy 373 TACCTTCCTTTTGGATGAGGAAACGAGTGTGATGTTGGAGCAAAAGGATCAATATT 432
Db 189 TACCTTCCTTTTGGATGAGGAAACGAGTGTGATGTTGGAGCAAAAGGATCAATATT 248
Qy 433 TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGATCTTACAC 492
Db 249 TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGATCTTACAC 308
Qy 493 CAGAAGAGATGAATGCAAGTGGCTGGAAAAAGACATCTCTGAAAGAAATGTGCTAATTTTCA 552
Db 309 CAGAAGAGATGAATGCAAGTGGCTGGAAAAAGACATCTCTGAAAGAAATGTGCTAATTTTCA 368
Qy 553 CAAGGTACTTAAGCATATATCAAGTCACTTGTGACCTGTGGAACGGGGGCTTTTCA 612
Db 369 CAAGGTACTTAAGCATATATCAAGTCACTTGTGACCTGTGGAACGGGGGCTTTTCA 428
Qy 613 TCCAAATTTGACCTTACATTTGAAATTTGGACATCATCTGAGGACAAATATTTTAAAGCTGA 672
Db 429 TCCAAATTTGACCTTACATTTGAAATTTGGACATCATCTGAGGACAAATATTTTAAAGCTGA 488
Qy 673 GAACTCAATTTTGAAGACCGCTGGGAGAGTCCATATGACCTTAAGCTGTGACAGC 732
Db 489 GAACTCAATTTTGAAGACCGCTGGGAGAGTCCATATGACCTTAAGCTGTGACAGC 548
Qy 733 ATCCCTTTTAATAGATGAGAAATTAATCTCTGGAATCTGACGTGATTTTATGGGGCGAGA 792
Db 549 ATCCCTTTTAATAGATGAGAAATTAATCTCTGGAATCTGACGTGATTTTATGGGGCGAGA 608
Qy 793 CTTTGTCTATCTTCGAACTCTTGGGACCAACCAATCAGGACAGGACGATGATTC 852
Db 609 CTTTGTCTATCTTCGAACTCTTGGGACCAACCAATCAGGACAGGACGATGATTC 668
Qy 853 CAGGTGGCTCAATGATCCAAAGTTTCAATGATGCGCCACCTCATCTCAGAGAGTGACAATCC 912
Db 669 CAGGTGGCTCAATGATCCAAAGTTTCAATGATGCGCCACCTCATCTCAGAGAGTGACAATCC 728
Qy 913 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAAAATGCAATAGATGAGAACTCTGG 972
Db 729 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAAAATGCAATAGATGAGAACTCTGG 788
Qy 973 AAAAGCTACTCACGCTAGATAGGTCAGATATGAAAGATGACCTTTGGAGGACAGAG 1032
Db 789 AAAAGCTACTCACGCTAGATAGGTCAGATATGAAAGATGACCTTTGGAGGACAGAG 848

REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:

INFORMATION FOR SEQ ID NO: 53:

SEQUENCE CHARACTERISTICS:

LENGTH: 2601 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cdna

FEATURE:

NAME/KEY: CDS

LOCATION: 16..2331

US-09-060-610-53

Query Match 92.6%; Score 2508.4; DB 4; Length 2601;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	193	CTGCAGCATGGGCTGTTAACTAGGATTTGCTGCTCTTTCTGGGAGTATTACTTACAGC	252
DB	9	CTGCAGCATGGGCTGTTAACTAGGATTTGCTGCTCTTTCTGGGAGTATTACTTACAGC	68
QY	253	AAGAGCAAACTATCAGATGGGAAGAACAAATGTCGCAAGGCTGAAATATCTTACAAAGA	312
DB	69	AAGAGCAAACTATCAGATGGGAAGAACAAATGTCGCAAGGCTGAAATATCTTACAAAGA	128
QY	313	AATGTTGGAATCAACAATGTGATCACTTTCAATGGCTTGGCCACAGCTCCAGTTATCA	372
DB	129	AATGTTGGAATCAACAATGTGATCACTTTCAATGGCTTGGCCACAGCTCCAGTTATCA	188
QY	373	TACCTTCCCTTTGGATGAGAAACGAGTAGGCTGTATGTTGGAGCAAAAGGATCAATAT	432
DB	189	TACCTTCCCTTTGGATGAGAAACGAGTAGGCTGTATGTTGGAGCAAAAGGATCAATAT	248
QY	433	TTCAATTCGACCTGGTTAAATCAAGGATTTTCAAAAGATTTGTGCGCAGTATCTTACAC	492
DB	249	TTCAATTCGACCTGGTTAAATCAAGGATTTTCAAAAGATTTGTGCGCAGTATCTTACAC	308
QY	493	CAGAAGAGATGAATCAAGTGGGCTGGAAAAGACATCTCTGAAAGAAATGTGCTAAATTCAT	552
DB	309	CAGAAGAGATGAATCAAGTGGGCTGGAAAAGACATCTCTGAAAGAAATGTGCTAAATTCAT	368
QY	553	CAAGGTACTTAAGGCATATAATCAGACTCACTTTGACGCTGTGAAACGGGGGCTTTTCA	612
DB	369	CAAGGTACTTAAGGCATATAATCAGACTCACTTTGACGCTGTGAAACGGGGGCTTTTCA	428
QY	613	TCCAAATTCACCTACATTTGAAATGGACATCATCTGAGGACAAATATTTTAAAGCTGGA	672
DB	429	TCCAAATTCACCTACATTTGAAATGGACATCATCTGAGGACAAATATTTTAAAGCTGGA	488
QY	673	GAACTCACATTTTGAACCGCGCTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGC	732
DB	489	GAACTCACATTTTGAACCGCGCTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGC	548
QY	733	ATCCCTTTTAATAGATGGAGAAATATACTCTGGAACCTGAGCTGATTTTATGGGGCGAGA	792
DB	549	ATCCCTTTTAATAGATGGAGAAATATACTCTGGAACCTGAGCTGATTTTATGGGGCGAGA	608
QY	793	CTTTTGTATCTTCGGAACCTTTGGGCACCAACCAATCAGGACAGACGACATGATTC	852
DB	609	CTTTTGTATCTTCGGAACCTTTGGGCACCAACCAATCAGGACAGACGACATGATTC	668
QY	853	CAGGTGGCTCAATGATCCAAAGTTCAATAGTCCCACTCATCTCTCAGAGAGTGACAATCC	912
DB	669	CAGGTGGCTCAATGATCCAAAGTTCAATAGTCCCACTCATCTCTCAGAGAGTGACAATCC	728
QY	913	TGAAGATGACAAAGTATATCTTTTCTTCGCGTGAATGAATAGATGGAGAACATCTGG	972
DB	729	TGAAGATGACAAAGTATATCTTTTCTTCGCGTGAATGAATAGATGGAGAACATCTCTGG	788

QY	973	AAAAGCTACTCAGCTAGATAGGTGACATATGCAAGAAATGACTTTGGAGGGCACAGAAG	1032
DB	789	AAAAGCTACTCAGCTAGATAGGTGACATATGCAAGAAATGACTTTGGAGGGCACAGAAG	848
QY	1033	TCTGGTGAATAAATGGACAACATTCCTCAAGCTGCTGCTGATTTGCTCAGTGCCAGGTCC	1092
DB	849	TCTGGTGAATAAATGGACAACATTCCTCAAGCTGCTGCTGATTTGCTCAGTGCCAGGTCC	908
QY	1093	AAATGGCATTCACACTCATTTTGTATGAACTCGAGATGATATCTTAATGAATTTAAGA	1152
DB	909	AAATGGCATTCACACTCATTTTGTATGAACTCGAGATGATATCTTAATGAATTTAAGA	968
QY	1153	TCCTAAAAAATCCAGTTGTATATGAGTGTTCAGCACTTCCAGTAACATTTTCAAGGGATC	1212
DB	969	TCCTAAAAAATCCAGTTGTATATGAGTGTTCAGCACTTCCAGTAACATTTTCAAGGGATC	1028
QY	1213	AGCCGTGTGTATGTATAGCATGATGTCGAGAAAGGGTGTTCCTTGGTCCATATGCCCA	1272
DB	1029	AGCCGTGTGTATGTATAGCATGATGTCGAGAAAGGGTGTTCCTTGGTCCATATGCCCA	1088
QY	1273	CAGGATGGACCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCAGGCC	1332
DB	1089	CAGGATGGACCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCAGGCC	1148
QY	1333	AGGAACCTTGTCCAGCAAAACATTTGGTGGTGTTCAGTCTTCAAAAGGACCTTCTCTGATGA	1392
DB	1149	AGGAACCTTGTCCAGCAAAACATTTGGTGGTGTTCAGTCTTCAAAAGGACCTTCTCTGATGA	1208
QY	1393	TGTTATAACCTTTGCAAGAGTCAATCCAGCATGTACAAATCCAGTGTTCCTATGAACAA	1452
DB	1209	TGTTATAACCTTTGCAAGAGTCAATCCAGCATGTACAAATCCAGTGTTCCTATGAACAA	1268
QY	1453	TCGCCAATAGTATCAAAACGGATGTAATATCAATTTACAAATTTGCTGTAGACCG	1512
DB	1269	TCGCCAATAGTATCAAAACGGATGTAATATCAATTTACAAATTTGCTGTAGACCG	1328
QY	1513	AGTGATGCAAGATGGACATGATGTTATGTTTATCGGAACAGATGTTGGACCGT	1572
DB	1329	AGTGATGCAAGATGGACATGATGTTATGTTTATCGGAACAGATGTTGGACCGT	1388
QY	1573	TCTTAAAGTAGTTCATTTCTTAAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGA	1632
DB	1389	TCTTAAAGTAGTTCATTTCTTAAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGA	1448
QY	1633	AGAAATGACAGTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTTCCACTAAGCA	1692
DB	1449	AGAAATGACAGTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTTCCACTAAGCA	1508
QY	1693	GCAACAACTATATTTGGTTCAACGGCTGGGGTTGCCAGCTCCCTTTTACACCGGTGTA	1752
DB	1509	GCAACAACTATATTTGGTTCAACGGCTGGGGTTGCCAGCTCCCTTTTACACCGGTGTA	1568
QY	1753	TATTTACGGGAACCGTGTGCTGAGTGTGCTCGCCGAGACCTTTACTGTGCTGGGA	1812
DB	1569	TATTTACGGGAACCGTGTGCTGAGTGTGCTCGCCGAGACCTTTACTGTGCTGGGA	1628
QY	1813	TGGTTCTGCTGTTCTGCTATTTTCCCACTGCAAGAGACGACACAGACATAT	1872
DB	1629	TGGTTCTGCTGTTCTGCTATTTTCCCACTGCAAGAGACGACACAGACATAT	1688
QY	1873	AAGAAATGGAGACCCACTGACTCTGTTTACAGCTTACACCATGATATACCATGGCCA	1932
DB	1689	AAGAAATGGAGACCCACTGACTCTGTTTACAGCTTACACCATGATATACCATGGCCA	1748
QY	1933	CAGCCCTGAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGACATTTTTGGAAATGAG	1992
DB	1749	CAGCCCTGAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGACATTTTTGGAAATGAG	1808
QY	1993	TCCGAAGTCCGAGAGAGGGCTGTTTATTTGCAATTTCCGAGGCGCAATGAGAGCGGAA	2052
DB	1809	TCCGAAGTCCGAGAGAGGGCTGTTTATTTGCAATTTCCGAGGCGCAATGAGAGCGGAA	1868

QY 2053 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGTAG 2112
Db 1869 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGTAG 1928
QY 2113 TCTACAAACAGAGATTCAAGGCAATTTACCTCTGCCATGCGTGGACATGGGTTTATACA 2172
Db 1929 TCTACAAACAGAGATTCAAGGCAATTTACCTCTGCCATGCGTGGACATGGGTTTATACA 1988
QY 2173 AACTCTTCTTAAGTTAAAGTAACTGGAAGTCAATTGACACAGAGCAATTTGGAAGAACTTCTTCA 2232
Db 1989 AACTCTTCTTAAGTTAAAGTAACTGGAAGTCAATTGACACAGAGCAATTTGGAAGAACTTCTTCA 2048
QY 2233 TAAAGATGATGGAGTGGCTCTTAAGACCAAGAAATGTCATAGATGACACCTAG 2292
Db 2049 TAAAGATGATGGAGTGGCTCTTAAGACCAAGAAATGTCATAGATGACACCTAG 2108
QY 2293 CCAGAGAGTCTGGTACAGAGACTTCATGAGCTCATCAACACCCCAATCTCAACAGCAT 2352
Db 2109 CCAGAGAGTCTGGTACAGAGACTTCATGAGCTCATCAACACCCCAATCTCAACAGCAT 2168
QY 2353 GGATGAGTCTGTGAACAAGTTTGGAAAGAGGACCGAAACAAACGTCGGCAAGGCCAGG 2412
Db 2169 GGATGAGTCTGTGAACAAGTTTGGAAAGAGGACCGAAACAAACGTCGGCAAGGCCAGG 2228
QY 2413 ACATACCCAGGACAGTAAACAATGGAAGCACTTACAGAAATAGAAAGGTAGAAA 2472
Db 2229 ACATACCCAGGACAGTAAACAATGGAAGCACTTACAGAAATAGAAAGGTAGAAA 2288
QY 2473 CAGGAGGACCCAGCAATTTGAGAGGCGACCCAGGAGTGTCTGAGCTGCAATACCTTACA 2532
Db 2289 CAGGAGGACCCAGCAATTTGAGAGGCGACCCAGGAGTGTCTGAGCTGCAATACCTTACA 2348
QY 2533 AACCTCAAAACAAGTAGAACTTGCCTAGACAATAAAGTGGAAAAACAATGCAATATACAT 2592
Db 2349 AACCTCAAAACAAGTAGAACTTGCCTAGACAATAAAGTGGAAAAACAATGCAATATACAT 2408
QY 2593 GAACCTTTTTCATGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2652
Db 2409 GAACCTTTTTCATGCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2468
QY 2653 CCAATTATATAATTAATCCATGAGTAACTTTCCTTAATAGGCTTTTTCCTCC 2702
Db 2469 CCAATTATATAATTAATCCATGAGTAACTTTCCTTAATAGGCTTTTTCCTCC 2518

RESULT 6

PCT-US94-10151A-53

; Sequence 53, Application PC/TUS9410151A

; GENERAL INFORMATION:

; APPLICANT: The Regents of the University of California

; TITLE OF INVENTION: The Semaphorin Gene Family

; NUMBER OF SEQUENCES: 66

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: FLEHR HOEBACH TEST ALBRITTON & HERBERT

; STREET: 4 Embarcadero Center, Suite 3400

; CITY: San Francisco

; STATE: CA

; COUNTRY: USA

; ZIP: 94111-4187

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US94/10151A

; FILING DATE: 13-SEP-1994

; CLASSIFICATION:

; ATTORNEY/AGENT INFORMATION:

; NAME: Osman, Richard A.

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: FP-58750-PC/RAO

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 781-1989

; TELEFAX: (415) 398-3249

; TELE: 910 277299 FHT UR

; INFORMATION FOR SEQ ID NO: 53:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2601 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 16...2331

PCT-US94-10151A-53

Query Match

Best Local Similarity 92.6%; Score 2508.4; DB 5; Length 2601;

Matches 2509; Conservative 0; Pred. No. 0;

Mismatches 1; Indels 0; Gaps 0;

QY 193 CTGACGATGGGCTGGTTAACTAGGATTTGTCTGCTTTTCTGGGAGTATTACTTTACAGC 252
Db 9 CTGACGATGGGCTGGTTAACTAGGATTTGTCTGCTTTTCTGGGAGTATTACTTTACAGC 68
QY 253 AAGAGCAAACTATCAGAAATGGGAAAGAAACAATGTGCCAAGGCTGAAAATTTATCCTACAAAGA 312
Db 69 AAGAGCAAACTATCAGAAATGGGAAAGAAACAATGTGCCAAGGCTGAAAATTTATCCTACAAAGA 128
QY 313 AATGTGGAAATCCCAACCAATGTGATCACTTTTCAATGGCTTGGCCAAAGCTCCAGTTATCA 372
Db 129 AATGTGGAAATCCCAACCAATGTGATCACTTTTCAATGGCTTGGCCAAAGCTCCAGTTATCA 188
QY 373 TACCTTCTCTTTGGATGAGGAAACGAGTAGGCTGTATGTTGGAGCAAGGATCACATATT 432
Db 189 TACCTTCTCTTTGGATGAGGAAACGAGTAGGCTGTATGTTGGAGCAAGGATCACATATT 248
QY 433 TTCAATTCGACCTGTTTAATATCAAGGATTTTCAAAAGATTTGTGTGGCCAGTATCTTACAC 492
Db 249 TTCAATTCGACCTGTTTAATATCAAGGATTTTCAAAAGATTTGTGTGGCCAGTATCTTACAC 308
QY 493 CAGAAGAGATGAATGCAAGTGGGCTGGAAAGACATCTCTGAAAGAAATGCTGCTAATTTTCAT 552
Db 309 CAGAAGAGATGAATGCAAGTGGGCTGGAAAGACATCTCTGAAAGAAATGCTGCTAATTTTCAT 368
QY 553 CAAGGTACTTAAAGGCATATATATCAGACTCACTTGTAGCCCTGTGGAACGGGGCTTTTCA 612
Db 369 CAAGGTACTTAAAGGCATATATATCAGACTCACTTGTAGCCCTGTGGAACGGGGCTTTTCA 428
QY 613 TCCAAATTTGCACTTACATTTGAAATTTGGAATTTGGAATTTGGAATTTTAAAGCTGGA 672
Db 429 TCCAAATTTGCACTTACATTTGAAATTTGGAATTTGGAATTTTAAAGCTGGA 488
QY 673 GAACCTACATTTTGAAGACCGGCTGGGAGAGTCCATATGACCTAAGCTGCTGACAGC 732
Db 489 GAACCTACATTTTGAAGACCGGCTGGGAGAGTCCATATGACCTAAGCTGCTGACAGC 548
QY 733 ATCCCTTTTAAATAGATGGAGAAATTTATATCTCTGGAATCTGCAAGCTGATTTTATGGGGCCAGA 792
Db 549 ATCCCTTTTAAATAGATGGAGAAATTTATATCTCTGGAATCTGCAAGCTGATTTTATGGGGCCAGA 608
QY 793 CTTTGTCTATCTTCGAACTCTTTGGGCAACCCCAATCAGGACAGAGCAAGATGATTC 852
Db 609 CTTTGTCTATCTTCGAACTCTTTGGGCAACCCCAATCAGGACAGAGCAAGATGATTC 668
QY 853 CAGGTGGCTCAATGATCCAAAGTTTCAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 912
Db 669 CAGGTGGCTCAATGATCCAAAGTTTCAATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 728
QY 913 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGAGAGCAACTCTGG 972
Db 729 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGAGAGCAACTCTGG 788
QY 973 AAAAGCTACTCAGCTAGAAATAGGTTCAGATATGCAAGAAATGATTTTGGAGGACAGAAAG 1032
Db 973 AAAAGCTACTCAGCTAGAAATAGGTTCAGATATGCAAGAAATGATTTTGGAGGACAGAAAG 1032

TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1481 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: both
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: CDS
LOCATION: 50..1480
US-08-136-922-1

Query Match 52.2%; Score 1415.2; DB 1; Length 1481;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1420; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 899 GAGAGTGAACCTCTGGAAGATGACAAAGTATACCTTTCTCCGTGAAATGCAATAGAT 958
DB |||||
53 GAACATGACAAATCCTGAAGATGACAAAGTATACCTTTCTCCGTGAAATGCAATAGAT 112
QY 959 GGAGAACACTCTGGAAAGCTACTCAGCTAGATAGTTCAGATGCAAGATGCACTTT 1018
DB |||||
113 GGAGAACACTCTGGAAAGCTACTCAGCTAGATAGTTCAGATGCAAGATGCACTTT 172
QY 1019 GGAGGGCACAGAGTCTGGTGAATAAATGGACAAATTCCTCAAGCTCGTCTGATTTGC 1078
DB |||||
173 GGAGGGCACAGAGTCTGGTGAATAAATGGACAAATTCCTCAAGCTCGTCTGATTTGC 232
QY 1079 TCAGTGCAGGTCCTAAATGGCAATGACACTATTTTGATGAATCAGAGATGATTTCTTA 1138
DB |||||
233 TCAGTGCAGGTCCTAAATGGCAATGACACTATTTTGATGAATCAGAGATGATTTCTTA 292
QY 1139 ATGAACTTTAAAGATCCTAAATCCAGTTGATATGAGTGTGATGAGAGGGTGTCTCTT 1258
DB |||||
293 ATGAACTTTAAAGATCCTAAATCCAGTTGATATGAGTGTGATGAGAGGGTGTCTCTT 352
QY 1199 ATTTTCAAGGGATCAGCCGTGTGTATGTATAGCATGATGTGAGAGGGTGTCTCTT 1258
DB |||||
353 ATTTTCAAGGGATCAGCCGTGTGTATGTATAGCATGATGTGAGAGGGTGTCTCTT 412
QY 1259 GGTCCATATGCCACAGGATGGACCCCAATATCAATGATGGTGCCTTTATCAAGGAAGTCT 1318
DB |||||
413 GGTCCATATGCCACAGGATGGACCCCAATATCAATGATGGTGCCTTTATCAAGGAAGTCT 472
QY 1319 CCCTATCCAGCCGAGGACTTGTCCAGCAAAATTTGGTGGTGGTGGTGGTGGTGGTGGTGG 1378
DB |||||
473 CCCTATCCAGCCGAGGACTTGTCCAGCAAAATTTGGTGGTGGTGGTGGTGGTGGTGGTGG 532
QY 1379 GACCTTCTGATGATGTTATACCTTTGCAAGAGTCAATCCAGCCATGTACAAATCCAGTG 1438
DB |||||
533 GACCTTCTGATGATGTTATACCTTTGCAAGAGTCAATCCAGCCATGTACAAATCCAGTG 592
QY 1439 TTTCCTATGAACAATCGCCCAATAGTATGATCAAAACGGATGTAATTAATTAATTAACAAA 1498
DB |||||
593 TTTCCTATGAACAATCGCCCAATAGTATGATCAAAACGGATGTAATTAATTAATTAACAAA 652
QY 1499 ATTTGCTGACCGAGTGGATGACAGATGACAGATGATGATGATGATGATGATGATGATGATG 1558
DB |||||
653 ATTCGCTGACCGAGTGGATGACAGATGACAGATGATGATGATGATGATGATGATGATGATG 712
QY 1559 GATGTTGGGACCGTTCTTAAAGTATGTTCAATTTCTAAGGAGACTTGGTATGATTTAGAA 1618
DB |||||
713 GATGTTGGGACCGTTCTTAAAGTATGTTCAATTTCTAAGGAGACTTGGTATGATTTAGAA 772
QY 1619 GAGGTTCTGCTGGAAGAAATGACAGTTTTCGGGAACCGACTGCTATTTTCAGCAATGAG 1678
DB |||||
773 GAGGTTCTGCTGGAAGAAATGACAGTTTTCGGGAACCGACTGCTATTTTCAGCAATGAG 832
QY 1679 CTTTCCACTAACGAGCAACTATATATTTGTTTCAAGCGTGGGGTGGCCAGCTCCCT 1738
DB |||||

DB 833 CTTTCCACTAAGCAGCAACAACTATATATTTGGTTCAACGGCTGGGGTTGCCAGCTCCCT 892
QY 1739 TTACACCGGTGTGATATTTTACGGGAAACGGTGTGTGCTGAGTGTGTGCTCGCCGAGACCT 1798
DB |||||
893 TTACACCGGTGTGATATTTTACGGGAAACGGTGTGTGCTGAGTGTGTGCTCGCCGAGACCT 952
QY 1799 TACTGTGCTGGGATGGTTCGCAATGTTCTCGCTATTTTCCACTGCAAAAGAGACGACA 1858
DB |||||
953 TACTGTGCTGGGATGGTTCGCAATGTTCTCGCTATTTTCCACTGCAAAAGAGACGACA 1012
QY 1859 AGACGACAAGATATAAGAAATGGAGACCACTGCTCACTCTTTCAGACTTACACCATGAT 1918
DB |||||
1013 AGACGACAAGATATAAGAAATGGAGACCACTGCTCACTCTTTCAGACTTACACCATGAT 1072
QY 1919 AATCACCATGGCCACAGCCCTGAAAGAGAGAAATCATCTATGTGTAGAGAAATAGTAGCACA 1978
DB |||||
1073 AATCACCATGGCCACAGCCCTGAAAGAGAGAAATCATCTATGTGTAGAGAAATAGTAGCACA 1132
QY 1979 TTTTGGAAATGCAATCGAAGTTCGAGAGAGCGGTGTCTATTTGGCAATTTCCAGAGGCGA 2038
DB |||||
1133 TTTTGGAAATGCAATCGAAGTTCGAGAGAGCGGTGTCTATTTGGCAATTTCCAGAGGCGA 1192
QY 2039 AATGAAGAGCCAAAGAGAGATCAGAGTGTATCATATCATCATCATCATCATCATCATCAT 2098
DB |||||
1193 AATGAAGAGCCAAAGAGAGATCAGAGTGTATCATATCATCATCATCATCATCATCATCAT 1252
QY 2099 CTTCTGTACCTAGTCTTACAAAGAGATTTTCAAGCAATTTACCTCTGCGCGGTGGA 2158
DB |||||
1253 CTTCTGTACCTAGTCTTACAAAGAGATTTTCAAGCAATTTACCTCTGCGCGGTGGA 1312
QY 2159 CATGGTTTCATCAAACTCTTTTAAAGTAAACCTTGAAGTCAATTTGACACAGAGCATTTG 2218
DB |||||
1313 CATGGTTTCATCAAACTCTTTTAAAGTAAACCTTGAAGTCAATTTGACACAGAGCATTTG 1372
QY 2219 GAAAGAACTTCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2278
DB |||||
1373 GAAAGAACTTCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2278
QY 2279 AGCATGACACCTAGCCAGAGGTCTGGTACAGAGACTTTCATGCGAGCTC 2326
DB |||||
1433 AGCATGACACCTAGCCAGAGGTCTGGTACAGAGACTTTCATGCGAGCTC 1480

RESULT 8

US-09-308-179B-2
; Sequence 2, Application US/09308179B
; Patent No. 6436669
; GENERAL INFORMATION:
; APPLICANT: INAGAKI, Shinobu
; APPLICANT: FURUYAMA, Tatsuo
; TITLE OF INVENTION: NOVEL SEMAPHORIN GENES (1)
; FILE REFERENCE: 0020-4562p
; CURRENT APPLICATION NUMBER: US/09/308,179B
; CURRENT FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/JP97/04111
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: JAPAN 321068/1996
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2898
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2898)
; OTHER INFORMATION: Strandedness: Double-stranded
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(2898)
; OTHER INFORMATION: any n = a, c, g, t, unknown, or other
US-09-308-179B-2

Query Match	22.0%;	Score 596.4;	DB 4;	Length 2898;
Best Local Similarity	56.3%;	Pred. No. 1.1e-175;		
Matches 1229;	Conservative	0;	Mismatches 931;	Indels 24; Gaps 5;

QY	223	CTGTCTTTCTGGGAGTATTACTTACACAGAGCAAACTATCAGAATGGGAAGACAA	282
DB	396	CTTGGCTGCTCTGGGGTCACCTGCTGGAACTCTGGACCCCCAGGTCACTCCGGGAACCCCTC	455
QY	283	TGTGCCAAGCTGAAATATCTTACAAAGAAATGTTGGAATCCAAACATGTGATCACTTT	342
DB	456	CTACCCAGCTACGGCTGTACATNAGAACTTTTGGAACTGAATAGGACTTCAATATT	515
QY	343	CAATGGCTTGGCCAAACAGCTCCAGGTATTATCATCTTCTCTTTTGGATGAGGAACGGAGTAG	402
DB	516	TCAAAGCCCTTGGATTCTTGATCTCTCATACATGCTGCTGGATGAGTATCAAGAACG	575
QY	403	GCTGTATGTTGGAGCAAAAGATCACATATTTTCATTCGACCTGGTTAATATCA---AGGA	459
DB	576	GCTCTTTGTGGAGGCAGAGACCTTGTCTATTCCCTGAACTTGGAAACGAGTCAGTGACGG	635
QY	460	TTTTCAAAGATGTGTGCCAGTATCTTACACAGAGAGATGAATCAAGTCGGGCTGG	519
DB	636	CTACAGAGAGATATCTGGCCGAGCAGCAGCTAAGGTAGAGAATGCATAATGAAGG	695
QY	520	AAAAGACATCTTGAAAGAAATGCTAAATTTTCATCAAGGTACTTAAAGGCATATAATCAGAC	579
DB	696	ANAAGA---CGCAATGAGTGTGCCATATATATCCGGGTTTGGCATCACACAGGAC	752
QY	580	TCACTTGTACGCTGTGGAAACGGGGCTTTTCATCCAAATTTGCACTCATATTGAATTGG	639
DB	753	ACACCTTCTGACCTGTGTACTGGAGCTTTTGATCCACACTGTGCTTCATCAGATCGG	812
QY	640	ACATCATCTGAGGACAAATATTTTAAAGCTGAGAACTCACATTTTGAACGCCGCTGG	699
DB	813	GCACCAATCAGAGAAACCCCTGTTTCACTCGAGTCAACAGATCTGAGAGAGAAAGGG	872
QY	700	GAAAGTTCATATGACCCCTAAGCTGTGACAGCATCCCTTTTAAATAGATGAGAAATATA	759
DB	873	CAGATGCTCTTTGACCCCAACTCCTCTCTTTGTGTCACGCTAGTTGGGAATGAGCTTT	932
QY	760	CTCTGAACTGCACTGATTTTATGGGGGAGACTTTTGTCTATCTTCCGAATCTTTGGGCA	819
DB	933	TGCTGGACTCTACAGTGACTATTGGGGGAGAGACTCGGCGATCTTCCGACAGATGGGGAA	992
QY	820	CCACACCCCAATCAGACAGCAGCATGATTCACAGGTGGCTCAATGATCCAAAGTTTAT	879
DB	993	GTTAGGCCATATTCCGACCTGAGCATGACGATGAGGCGCTCTGAAAGAACCAAAATTTGT	1052
QY	880	TAGTGCCCACTCATCTCAGAGAGTGACAATCTCTCAAGATGACAAGATATATCTTTTCTT	939
DB	1053	AGGTTCAATATGATTCCTGTATGACAGACCGAGATGACACAAATATGATCTTTTCTT	1112
QY	940	CGGTGAAATGCAATAGATGGAGAACACTCTGTGAAAGCTACTCACGCTAGAAATAGTCA	999
DB	1113	TACTGAGAAGCGCTGGAGGCGAGAGAACAAAGCCACACAGCTATACACCCGAGTGGGGCG	1172
QY	1000	GATATCAAGATGACTTTTGGAGGCAAGAGTCTGTGATTAATTAATGGACAACTTCTCT	1059
DB	1173	GCTGTGCTGATGATCATGGAGGACAGAGAAATCTGTGTGAACAAAGTGGAGCACTTTCTCT	1232
QY	1060	CAAAGCTGCTGTGATTTGCTCAGTGCAGGTCCAAATGGCAATGACACTCAATTTTGATGA	1119
DB	1233	TAAAGCGCGCTGGTTGCTCAGTGCAGGGAATGAATGGAATGACACATATTTTGAGA	1292
QY	1120	ACTGCAAGATGATTTCTTAATGAACTTTTAAAGATCCTTAAAGATCCAGTTGATATGAGT	1179
DB	1293	ACTAGAGGATGTGTTTTTACTGCCGACAGAGATCCTAAGAAATCCAGTGATATTGGACT	1352
QY	1180	GTTTACGATCTCCAGTAACTTTTCAAGGGATCAGCGGTGTGATGTATGATGATGATGA	1239
DB	1353	GTTTAACTACAGCAATATATTAGAGGCCATGCTGTATGTGTGATCATCATGTCAAG	1412

QY	1240	TGTGAGAAGGGTGTCTTGTGTCATATATGCCCAACAGGATGGACCCCAACTATTCATCGGT	1239
DB	1413	TATCCGGGAAGCCTTTATGTGCCCATATGTCTATAAAAGAGGCCCTGAATACCACTGGTC	1472
QY	1300	GCCTTATCAAGGAGAGTCCCTATCCACGSCCAGGAACCTTGTCCCAAGCAAAACATTGG	1359
DB	1473	ACTATATGAAGGAAAGTCCCTTACCAGAGCCCTGGTTCTGTGCCAGCAAAAGTAAACGG	1532
QY	1360	TGG---TTTTGACTCTCAAAAGGACCTTCTGTGATGATGTTATATACCTTTTGCAGAAAGTCA	1416
DB	1533	AGGCAAGTATGGAACCAACCAAGATATACCCGATGACGCCATCCGGTTTCGCAAGGATGCA	1592
QY	1417	TCCAGCCATGTACAAATCCAGTGTTCCTATGAAACAATCGCCCAATAGTGATCAAAACGGA	1476
DB	1593	TCCTCTAATGTATCAGCCCAATAAACCTGTTCATAAAAAAACCAATATCTGGTAAAAACAGA	1652
QY	1477	TGTAAATTTATCAATTTACAAATTTGCTGTAGACCGAGTGATGTCAGAGATGGAACAGTA	1536
DB	1653	TGGAATAATCAACCTGTAGGCAACTTGCCTGGATTCGGGTGGAAGCGGAGGATGGCCAGTA	1712
QY	1537	TGATGTTATGTTTATCGGAACAGATGTTGGGACCGTTCTTAAAGTAGTGTTCATTTCTCTAA	1596
DB	1713	TGACGCTTATTTATTTGGGACAGACAGAGAAATTTGCTGAAAGTATCACAAATTTCACAA	1772
QY	1597	GGAGACTTGGTATGATTTAGAAAGGTTCTCTGGAAGAAATGACAGTGTTCGGGAACC	1656
DB	1773	CCAAGAAACAGAGTGGATGGAGGAAGTCATTTCTAGAGGAACCTTCAAAATATTCAGAGATCC	1832
QY	1657	GACTGCTATTTCAAGATGAGGCTTTTCCATTAAGCAGCAACAACATATATATTGGTTCAAC	1716
DB	1833	AGCCCTATCATTTTCTATGGAAATTTCTTCAAGAGACAACAGCTTTTACATTTGGATCAGC	1892
QY	1717	GGCTGGGCTGCCACAGCTCCCTTTACACCGGTGTGATATTTACGGGAAACGCTGTGCTGA	1776
DB	1893	CTCTGCTGTGGCAACAAGTCAGATTCATCTACGCAATGATGCGAGTCTGTGTGCTGA	1952
QY	1777	GTGTTGCTCCGCCGAGACCTTACTGTCTGGGATGGTGTCTGCAATGTTCTCGCTATTTT	1836
DB	1953	CTGCTGCTGCTCGAGACCCGTACTGTGCTGGATGGCATATCTCTGCTCCAGGTACTA	2012
QY	1837	TCCAC-----TGCAAAGAGACGCAACAGACGACAAGATATAAGAATGGAGNCCC	1887
DB	2013	CCCAACAGGTGCACACGCAAGAGAGAGGTTTCCGAGGACGAGCGTTCGGCATGGCAACGC	2072
QY	1888	ACTGACTCACTGTTTCAGACTTTACACCATGATTAATCACATGGCCACACGCTTGAAGAG	1947
DB	2073	CGCCCAACAGTGTCTTGGACAGCAATTTTGTGGAGACGCGTTGGACAGGACTGAGAGAG	2132
QY	1948	AATCATCTATGTTGTAGAGATAGTAGCAATTTTTTGGAAATGCAAGTCCGAAGTCGAGAG	2007
DB	2133	GCTGGCTTATGGCATAGAGACACAGTACTCTGTTGGAATGCACCCCGCATCACTTACA	2192
QY	2008	AGCGCTGTGCTATTGGCAATTCAGAGCGGCAATGAAAGAGCGGAAAGAGAGATCAGAGT	2067
DB	2193	AGCAAAAGTCTATCTGTGTTTACAGAAAGGACGCGACGTAAAGAAAAGAGAGGTGAAGAC	225

	Query Match	3.1%	Score 84;	DB 4;	Length 2433;
	Best Local Similarity	47.6%;	Pred. No. 1.4e-15;		
	Matches 357;	Conservative 0;	Mismatches 375;	Indels 18;	Gaps 3;
Qy	356	AAACGCTCCAGTTATCATACCTTCCCTTTTGGATGAGGAACGGAGTAGGCTGTATGTGGA	415		
Db	34	AACATCTCCAACTACTACGCGCCCTTCTGCTGAGCCAGGATGGAAGACGCTGTATGTGGGG	93		
Qy	416	GCAAAGGATCACATATTTTCA-----TTGACCTGGTTAAATACAAAGGATTTT	463		
Db	94	GCCCGAGAGGCCCTCTTTGCATTTAACAGCAACCTCAGCTTTTGCACGGCGGGGAGTAC	153		

464 CAAAGATTGTGGCCAGTATCTTACACAGAGAGATGAATGCAAGTGGGCTGAAAA 523
154 CRAAGACTGTGGAGTGCAGATGTGACAGGACGACGAGTCAAGGGCAG 213
524 GACATCCTGAAGAAATGTGTAATTTATCAACAGGTACTTAAGGCATATATCAGACTAC 583
214 GACCCAAAGCGTGACTGTCAAAACTACATCAAGATCTCTGCACTCAACAGCACCAC 273
584 TTGTAGCCTGTGGAGCGGGCTTTTCATCCAAATTTGACCTACATTCGAAATGGACAT 643
274 CTGCTCACCTGTGGACGGCGCTTCAAGCCCTGTGCTTACATTCACATAGCGAGC 333
644 CATCTCTGAGGACAATATTTTAAAGCTGGGAACTCACATTTTGAAAAAGCGCGTGGGAG 703
334 TTATCTTAGCCCA---AGATGAGCGGGTAAATGTCTTCTGGAGATGGCAAGGTCTAT 390
704 AGTCCATATGACCTTAAGCTGTGACAGCATCCCTTTTAAATAGATGAGAAATATATCTCT 763
391 TGTCCCTTTGACCCCAACTTCAAGTCCAGGGCTCTGGTGTGATGGTGAAGTGTACACT 450
764 GGAACCTGCAGCTGATTTTATGGGGGAGACTTTGCTATCTTCCGAACTCTTGGGACACC 823
451 GGAACAGTCAGTAGCTTCCAGGGAAACGACCCAGCCATTTCCCGGAGCCAGAGTTCCCGC 510
824 CACCCAAATCAGGACAGCAGCATGATTCAGGTGGCTCAATG---ATCCAAAGTTCAAT 880
511 CCACCAAGACTGAGAGCTCCCTCACTGGCTACAAGACCTGCTTGTGGCCCTGGCT 570
881 AGTGCCCACTCATCTCAGAGTGTGAACAGTCCCTGGAAGATGACAAAGTATATCTTTCTTC 940
571 ACGTCCCGGAGAGCGCTGGCAGCCCATAGGTGATGATGAAGATCTACTTCTTCTTC 630
941 CGTGAATGCAATAGATGGAGAACACTCTGGAAGCTACTCACGCTAGATAGTCTAG 1000
631 AGCGAGACGGGCGAGAGTTGAGTTCTTTGAGAACACCATCGTGTCCGAGTGGCCGA 690
1001 ATATGCAAGAAATGACTTTGGAGGGCAGAGTCTGGTGAATAAATGGACAACTTCCTC 1060
691 GTCTGTAAAGGCGATGAGGTGGAGAGCGGGTGTTCAGCAACGCTGGACCTCTTCTC 750
1061 AAGCTCTGTGATTTGCTCAGTGCCAGGT 1090
751 AAGGCTCAGCTCTGTGCTCCCGGCTGAT 780

RESULT 11

US-08-556-422A-1
; Sequence 1, Application US/08556422A
; Patent No. 6576754
; GENERAL INFORMATION:
; APPLICANT: FREEMAN, Kathryn T.
; APPLICANT: SCHULTZE, Joachim J.
; APPLICANT: BOUSSIORIS, Vassiliki
; APPLICANT: NADLER, Lee M.
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES
; FILE REFERENCE: DFN-00SCP2
; CURRENT APPLICATION NUMBER: US/08/556,422A
; CURRENT FILING DATE: 1995-11-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 4157
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (88)...(2673)
US-08-556-422A-1

Query Match 3.1%; Score 83.2; DB 4; Length 4157;
Best Local Similarity 49.0%; Pred. No. 3.5e-15;

Matches 619; Conservative 0; Mismatches 568; Indels 75; Gaps 12;

QY 352 GSCCAACAGCTCCAGTATCATACCTTCTTTTGGATGAGGAACGAGTAGGCTGTATCT 411
Db 219 CCAGACATCTTACAACACTCTCAGCTTCTGCTGAGCGGAGCAAGGACACCTTGTACAT 278
QY 412 TGGCAAAAGGATCACATATTTTCATTCGACCTCGTTAATAT---CAAGGATTTTCAAAA 468
Db 279 AGGTCCCGGAGGGGGCTTCGCTGTGAACGCACTCAACATCTCCGAGAACGACATGA 338
QY 469 GATTGTGGCCAGTATCTTACACAGAGAGATGAATCAAGTGGGCTGGAAAAGACAT 528
Db 339 GGTGTATTTGGAAGGTCTCAGAAGACAAAAAGCAAAATGTGCAGAAAAGGGGAAATCAA 398
QY 529 CCTGAAAGAAATGTCTAATTTTCATCAAGGTACTTAAAGGCATATATACAGACTCACTTGA 588
Db 339 ACAGACAGGTGCTCAACTACATCCGGGTGCTGAGCCACTCAGCGGCACCTTCCCTTTA 458
QY 589 CGCTGTGGAACGGGGGCTTTTCATTCCTCAATTTGCACTTACATTTGAAATTTGGAACATCATCC 648
Db 459 CGTGTGGGACCAACGCAATTCAGCGCGCTGTGACCACCTGAATTTAAATC-----512
QY 649 TGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTTGAACCGCCCTGGGAAGAGTCC 708
Db 513 -----CTTTAAGTTTCTGGGAAAAA-----TGAAGATGGCAAGGAAGATGTCC 557
QY 709 ATATGACCTTAAGCTGTGACAGCATCCCTTTTAAATAGATGAGAAATATATCTCTGGAAC 768
Db 558 CTTTGACCCAGCACACAGCTACACATCCGTGATGGTGTGATGAGAACTTTTATTCGGGAC 617
QY 769 TGCAGCTGATTTTATGGGGCGAGACTTTTGTCTATCTTCCGAACTCTTGGGCAACCAACCC 828
Db 618 GTCTGATTAATTTTGGGAAGTGAACCCATCATCTCCGAAATTCCT---TCCACAGTCC 674
QY 829 AATCAGGACAGCAGCAGCATGATTCAGGTGCTCAATGATCAAAAGTTCATTTAGTGCCA 888
Db 675 TCTGAGGACAG---AATATGCAATCCCTTGGCTGAAACGAGCTAGTTCTGCTGTTGCTGA 731
QY 889 CTTCTAT-----CTCAGAGAGTGACAATCTTGAAGATGACAAAGTATATCTTTCTT 939
Db 732 CGTATTCGAAAAAGCCAGACAGCCCGAGCGGAGGATGACAGGTCTACTTCTTCTT 791
QY 940 CCGTGAATAATGCAATAGATGGAGAACACTCTTGGAAAAAGCTACTCACGCTAGATAGGTCA 999
Db 792 CACGAGGTGCTGTGGAGTATGATTTGTGTCAGGGTGTGATCCCAACGATAGCAAG 851
QY 1000 GATATGCAAGAAATGACTTTTGGAGGCGACAGAAAGTCTGTGTGAATAAATGGACAACTTCT 1059
Db 852 AGTGTGCAAGGGGACCCAGGGCGGCTGAGGACCTTGCAGAAAGAAATGGACCTCTCTCT 911
QY 1060 CAAAGCTCGTCTGATTTGCTCAGTGCAGGTCCAAATGGCAATTCACACTCAATTTTGTGA 1119
Db 912 GAAAGCCGACTCATCTGCTCCCGCCAGA-----CAGGGCTTGGTCTTCAATGT 962
QY 1120 ACTGCAGGATGTATTTCTTAATGAACTTTTAAAGATCTTAAAAATCCAGTTGTATGGAGT 1179
Db 963 CTGCGGAGTCTCTTGTGCTCAGTCCCGGCTGAGGTGCTGTGTTCTATGCACT 1022
QY 1180 GTTTACGACTTCCAGTAACTTTTCAAGGGATACGCGGTGTGTATGTATAGATGAGTGA 1239
Db 1023 CTTCAACCCACAGCTGAACAACTGTGGGGCTGTGCGCAGTGTGCGCTACAACCTGTCCAC 1082
QY 1240 TGTGAGAAGGTGTTCC-----TTGGTCCATATGCCACAG--GGATGGACCCAA 1287
Db 1083 AGCGAGAGGTCTTCTCCCAACGGGAAGTACATGAGAGACCAAGTGGAGCAGTCCCA 1142
QY 1288 CTATCAATGGGTGCTTTATCAAGGAAGAGTCCCTTATCCAGCGGACGAACTTGTCCAG 1347
Db 1143 CACCAAGTGGGTGCTATATATGCGCGGTACCCAGCGGCTGGAGCGTGCATGA 1202
QY 1348 CAAACAACTTTGGTGGTTTGTACTCTACAAA-----GGACCTTCTGTATGTATTAAC 1401
Db 1203 CAGCGAGGACCGGCGCCCAACTACACAGCTCCTTGAATTTGGCAGACAGACGCTGA 1262


```

; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,268
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2854 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 451..2640
;
US-08-835-268-57

Query Match 2.6%; Score 69.8; DB 1; Length 2854;
Best Local Similarity 48.0%; Pred. No. 4.3e-11;
Matches 351; Conservative 0; Mismatches 347; Indels 33; Gaps 4;

QY 918 ATGACAAAGTATCTTTCTTCCTGGAAGGCGTATGATGCTCTGTTTATAGGAAGTATGATG 1732
DB 1566 GGACCGTTCTTTAAAGTAGTCTTCAATTCCTAGGAGACTTGGTATGATTTAGAGAGGTTTC 1625
DB 1733 GCAAAAGTATGATAAAGCTTTGAACTCTGCCTCTCTTTGATTTCATCTGATCTAGTAGTAG 1792
QY 1626 TGCTGGAAGAA 1636
DB 1793 TTGTAATAGAA 1803

RESULT 14
US-09-060-692-57
; Sequence 57, Application US/09060692
; Patent No. 5935865
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,692
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2854 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 451..2640
;
US-09-060-692-57

Query Match 2.6%; Score 69.8; DB 2; Length 2854;
Best Local Similarity 48.0%; Pred. No. 4.3e-11;
Matches 351; Conservative 0; Mismatches 347; Indels 33; Gaps 4;

QY 918 ATGACAAAGTATCTTTCTTCCTGGAAGGCGTATGATGCTCTGTTTATAGGAAGTATGATG 977
DB 1094 ATGATTTTATATCTTCTTCTTCCGAGAGACTGCTGTTGAGTACATCACTGCGGAAGG 1153
QY 978 CTACTCAGCTAGATAGGTGAGATATGCAAGATGACTTTGGAGGGCAGAGAAGTCTGG 1037
DB 1154 CTATCTATTCAAGAGTTGCCAGAGTCTGTAACATGACGAAGGGCGCCCTCATCGGGTG 1213
QY 1038 TGATTAATGGACAACATTCCTCAAGCTCGTCTGATTTGCTCAGTGCAGGTCCTCAATG 1097
DB 1214 GTGACAGATGGACTTCTTTTGAATCAGCTGAACTGTTCCGTCCTCGAGATATATC 1273
QY 1098 GCATTCA---CACTCATTTTGTAGTCACTCAGAGTGTATTCCTAATGAACCTTTAAAGATC 1154
DB 1274 CATTTTACTTCAATGAATTCAGTCAAGAGTGAACATCAATGAAGGAATTAATGTTGGTTC 1333
QY 1155 CTAAAAATCCAGTTGTATATGAGTGTGTTACGACTTCCAGTAACACATTTTCAAGGGATCAG 1214
DB 1334 AAGTGGAGAACTCATCTACGGTGTCTTCCAGACACCAAGTGAACCTCTATTGGTGGCTCTG 1393
QY 1215 CGTGTGTATGATAGCATGATGATGAGAGGGGTTCCTTGTGTCATATGCCACA 1274
DB 1394 CTGTTTGTGCTTTCAGTAAAGTCAATCTTGAAGTCAATTTGATGTCATTTAAAGAGC 1453
QY 1275 GGGATGGACCAACTATCAATGGTGCCTTATCAAGGA---AGAGTCCCTATCCAGCGC 1331
DB 1454 AGGAACAGATGAACCTCAACTGGTGGCAGTGCCAGCCCTTAAGTGCAGAACCAAGGC 1513
QY 1332 CAGGAACCTTGTCCCGACAAAACATTTGGTGGTGTGTTTGAACCTCTACAAAGGACCTTCTCTGATG 1391
DB 1514 CTGGACAATGTGTAATGACATGCTGTAACA-----CTTCTCTGATG 1552
QY 1392 ATGTTTAACTTTGCAAGAGTCAATCCAGCCATGTAATCCAGTGTTCCTATGATGAACA 1451
DB 1553 TGTGTCTCAATTTGTAAAGTCAATACATGATGATGAGGCGCCGTCAGCATTTTAA 1612
QY 1452 ATCGCCCAATAGTATCAAAACGAGTAAATTTATCAATTTTACACAAATTTGCTGAGACC 1511
DB 1613 CTCGGCAATTTCTCATTCGATCAGCTTACAGTACAGATTTTACAAAATAGCTGTTGATC 1672
QY 1512 GAGTGA-----TGCAAGATGGACATGATGATGTTTATGTTTATCGGAACAGATGTTG 1565
```


Db 1154 CTATCTATTCAAGAGTTGCCAGAGTCTGTAAACATGACAAAGCGCGCCCTCATCAGGGTG 1213
Qy 1038 TGAATAAATGGACAACATCTCTCAAGAGTGGTCTGATTTGCTCAGTCCAGGTCCTCAAAATG 1097
Db 1214 GTGACAGATGGAGTCTCTTTTGAATCAGCTGAACTGTTCCGTCCTCGAGATTATC 1273
Qy 1098 GCATTGA---CACTCATTTGATGAAGTCCAGAGTGTATCTCAATGAATCTTAAAGATC 1154
Db 1274 CATTTTACTTCAATGAATTTCAAGTCAACAGTGACATCAATGAAGAAATATGTTGGTGC 1333
Qy 1155 CTAATAATCCAGTTGATATGAGTGTTCACGACTTCCAGTAACTTTTCAAGGGATCAG 1214
Db 1334 AAGTGGAGAACTCATCTACGGTGTCTTCAGACACCAAGTGAATCTATTTGGTGGCTGTG 1393
Qy 1215 CCGTGTGATGATAGCATGAGTGTGAGAAAGGTGTTCTTGGTCCATATGCCACACA 1274
Db 1394 CTGTTTGTGCTTCAAGTCAAGTCAATCTTGAAGTCAATTTGATGTTGATGTTTAAAGAGC 1453
Qy 1275 GGGATGGACCAACTATCAATGAGTGTGCTTATCAAGGA---AGAGTCCCTTATCCAGGC 1331
Db 1454 AGAAACGATGAAGTCAAACTGGTGGCAGTGCCTTAAAGTCCAGAACCAAGGC 1513
Qy 1332 CAGGAACCTTGTCCAGCAAAACATTTGGTGGTGTGTTGACTCTACAAAGGACCTTCTGTATG 1391
Db 1514 CTGGACAAATGTGATGATGACAGTCTGTA---CTTCTGATG 1552
Qy 1392 ATGTTATAACCTTTCAAGAGTCAATCCAGCCATGTAATTAATTAATTAATTAATTAATTA 1451
Db 1553 TGTCTGTCAATTTTGAAGTCAATCACTGATGATGAGGCGGTCGAGCAGATTTTAA 1612
Qy 1452 ATCGCCCAATGATGATCAAAAGGATGTAATTAATTAATTAATTAATTAATTAATTAATTA 1511
Db 1613 CTCGGCAATTTCTCATTCGGATCAGCTTACAGTACAGATTTTAAAGTATGATGATG 1672
Qy 1512 GAGTGA-----TGCAAGAGTGAAGTGTATGATGTTTATCGGAACAGATGTTG 1565
Db 1673 ACAAGTCCGAACAGATGGAAGCGTATGATGTTTATAGGACTGATGATG 1732
Qy 1566 GGACCGTCTTAAAGTGTGTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAAGAGTTTC 1625
Db 1733 GCAAAGTGAATAAAGCTTTGAAGTCTGCTCTCTTTGATTCATCTGATCTGATGATG 1792
Qy 1626 TGCTGGAAGAA 1636
Db 1793 TTGTAATAGAA 1803

RESULT 15

US-08-833-391-57.
Sequence 57, Application US/08833391
Patent No. 6013781
GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Matthews, David
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSER: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,391

FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/121,713
FILING DATE: 13-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 57:
SEQUENCE CHARACTERISTICS:
LENGTH: 2854 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
FEATURE:
NAME/KEY: CDS
LOCATION: 451..2640
US-08-833-391-57

Query Match 2.6%; Score 69.8; DB 3; Length 2854;

Best Local Similarity 48.0%; Pred. No. 4.3e-11;

Matches 351; Conservative 0; Mismatches 347; Indels 33; Gaps 4;

Qy 918 ATGCAAAAGTATACCTTTTCTTCGTTGAAATGCAATAGATGGAGAACATCTCTGAAAAAG 977
Db 1094 ATGATTTTATATCTTCTTCTTCGAGAGACTGCTGTTGAGTACATCAACTCGGGAAGG 1153
Qy 978 CTACTCAGCTAGATAGTTCAGATATGCAAGAATGACTTTTGGAGGGGACAGAAAGTCTGG 1037
Db 1154 CTATCTATTCAAGAGTTCAGAGTCTGTAAACATGACAAAGCGGCGCTCATCAGGGTG 1213
Qy 1038 TGAATAAATGGACAACATTCCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCCTCAAAATG 1097
Db 1214 GTGACAGATGAGTCTCTTTTGAATCAGCTCTGAACTGTTCCGTCCTCGAGATTATC 1273
Qy 1098 GCATTGA---CACTCATTTTGAATGAGTGTGAGGATGATTCCTAATGAATCTTAAAGATC 1154
Db 1274 CATTTTACTTCAATGAATTTCAAGTCAACAGTGCATCATTTGAAGAAATATGTTGCTC 1333
Qy 1155 CTAATAATCCAGTTGATATGAGTGTTCAGACTTCCAGTAAACATTTTCAAGGGATCAG 1214
Db 1334 AAGTGGAGAACTCATCTACGGTGTCTTCAGACACCAAGTGAATCTATTTGGTGGCTGTG 1393
Qy 1215 CCGTGTGATGATAGCATGAGTGTGAGAGGGTGTTCCTTGGTCCATATGCCACACA 1274
Db 1394 CTGTTTGTGCTTCCAGTATGAAGTCAATCTTGAAGTCAATTTGATGTTGCTCAATTTAAAGAGC 1453
Qy 1275 GGGATGGACCAACTATCAATCGGTGCTTATCAAGGA---AGAGTCCCTTATCCAGGC 1331
Db 1454 AGAAACGATGAAGTCAAACTGTTGGCAGTGCCTTAAAGTTCAGAACCAAGGC 1513
Qy 1332 CAGGAACCTTGTCCAGCAAAACATTTGGTGGTGTGTTGACTCTACAAAGGACCTTCTGTATG 1391
Db 1514 CTGGACAAATGTGATGATGACAGTCTGTA---CTTCTGATG 1552
Qy 1392 ATGTTATAACCTTTGCAAGAGTCAATCCAGCCATGTAATTAATTAATTAATTAATTAATTA 1451
Db 1553 TGTCTGTCAATTTTGAAGTCAATCACTGATGATGAGGCGGTCGAGCAGATTTTAA 1612
Qy 1452 ATCGCCCAATGATGATCAAAAGGATGTAATTAATTAATTAATTAATTAATTAATTAATTA 1511
Db 1613 CTCGGCAATTTCTCATTCGGATCAGCTTACAGTACAGATTTTAAAGTATGATGATG 1672
Qy 1512 GAGTGA-----TGCAAGAGTGAAGTGTATGATGTTTATCGGAACAGATGTTG 1565
Db 1673 ACAAGTCCGAACAGATGGAAGCGTATGATGTTTATAGGACTGATGATG 1732

Qy	1566	GGACCGTCTCTTAAAGTAGTTTCAATTCTTAAGGAGACTTGGTATGATTTTGAAGAGGTTTC	1625
Db	1733	GCAAACTGATAAAAGCTTTGAACCTCTGCTCTTTCCTTTCGATTCACTGATAGTAGTG	1792
Qy	1626	TGCTGGAAGAA	1636
Db	1793	TTGTAATAGAA	1803

Search completed: July 31, 2003, 17:20:36
Job time : 188 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 09:54:25 ; Search.time 129 Seconds
(without alignments)
10838.209 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 4896

Sequence: 1 aatctttatttatcatgatg.....agggttttttctctaataacc 2709

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1661050

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlh
-O=/cgn2_1/USPTO_spo1/US09774490/runat_03082003_095416_25331/app_query.fasta_1.2887
-DB=SPTREMBL_23 -QPT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09774490 @CGN 1 1 0 @runat_03082003_095416_25331 -NCPU=3
-NO_MMAP -LARGQUERY -NEG SCORE=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTREMBL_23:*
1: sp archaea:*
2: sp bacteria:*
3: sp fungi:*
4: sp human:*
5: sp invertebrate:*
6: sp mammal:*
7: sp mhc:*
8: sp organelle:*
9: sp phage:*
10: sp plant:*
11: sp rodent:*
12: sp virus:*
13: sp vertebrate:*
14: sp unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	3609.5	73.7	774	13 Q8JIW9	Q8JIW9 xenopus lae

2	2258	46.1	777	11	Q8BH34	Q8bh34 mus musculu
3	2252	46.0	777	11	Q8BMF6	Q8bmf6 mus musculu
4	2106.5	43.0	748	4	Q8TB71	Q8tb71 homo sapien
5	2102.5	42.9	754	4	Q8TDV7	Q8tdv7 homo sapien
6	2030	41.5	775	11	Q8QX23	Q8qx23 mus musculu
7	2005.5	41.0	756	13	Q8QG99	Q8qgu9 gallus gall
8	1905.5	38.9	635	4	Q8GXA0	Q8gxo0 homo sapien
9	1830	37.4	782	4	Q8NS98	Q8ns98 homo sapien
10	1238.5	25.3	416	4	Q8NCQ1	Q8ncq1 homo sapien
11	1176.5	24.0	457	4	Q8HBR1	Q8hbr1 homo sapien
12	1139.5	23.3	403	11	Q8BKQ6	Q8bkq6 mus musculu
13	1080	22.1	296	11	Q8JT29	Q8jj29 rattus norv
14	976.5	19.9	799	11	Q8BJC1	Q8bjc1 mus musculu
15	943	19.3	963	4	Q8C0C4	Q8c0c4 homo sapien
16	889	18.2	893	4	Q8C0B8	Q8c0b8 homo sapien
17	878.5	17.9	823	11	Q8BIR6	Q8bir6 mus musculu
18	854	17.4	284	11	O54948	O54948 mus musculu
19	854	17.4	1049	4	Q8P2H9	Q8p2h9 homo sapien
20	852.5	17.4	1030	4	Q8H2B6	Q8h2e6 homo sapien
21	836	17.1	550	11	Q8BIC3	Q8bic3 mus musculu
22	830.5	17.0	1073	4	Q8NFY4	Q8nfy4 homo sapien
23	828.5	16.9	587	11	Q8BUT0	Q8but0 mus musculu
24	828.5	16.9	605	11	Q8BXZ7	Q8bxz7 mus musculu
25	827.5	16.9	1017	4	Q8NFY5	Q8nfy5 homo sapien
26	826	16.9	595	11	Q8BKQ8	Q8bkq8 mus musculu
27	825.5	16.9	761	4	Q8WUA9	Q8wua9 homo sapien
28	822.5	16.8	998	4	Q8NFY6	Q8nfy6 homo sapien
29	812	16.6	1011	4	Q8NFY3	Q8nfy3 homo sapien
30	812	16.6	1022	4	Q8P249	Q8p249 homo sapien
31	793.5	16.2	1005	11	Q8EQ71	Q8eq71 mus musculu
32	791.5	16.2	697	5	Q8MLF1	Q8mlf1 drosophila
33	789.5	16.1	724	5	Q8V7Q7	Q8v7q7 drosophila
34	778.5	15.9	1202	4	Q8P283	Q8p283 homo sapien
35	762.5	15.6	935	4	Q86JF8	Q86jfb homo sapien
36	739.5	15.1	963	11	Q8IY36	Q8iy36 mus musculu
37	733.5	15.0	844	11	Q8BX08	Q8bx08 mus musculu
38	712	14.5	687	4	Q8BXH8	Q8bxh8 homo sapien
39	708	14.5	476	4	Q8NFY7	Q8nfy7 homo sapien
40	707	14.4	418	4	Q8GJT9	Q8gju9 homo sapien
41	695	14.2	770	5	O44253	O44253 drosophila
42	695	14.2	770	5	Q8V3M4	Q8v3m4 drosophila
43	666.5	13.6	367	4	Q8HAH9	Q8nah9 homo sapien
44	618	12.6	616	5	Q8V7P8	Q8v7p8 drosophila
45	602.5	12.3	920	11	Q8R4U4	Q8r4u4 rattus norv

ALIGNMENTS

RESULT 1

ID	Q8JIW9	PRELIMINARY;	PRT;	774 AA.
AC	Q8JIW9;			
DT	01-OCT-2002 (TRENBLrel. 22, Created)			
DT	01-OCT-2002 (TRENBLrel. 22, Last sequence update)			
DT	01-MAR-2003 (TRENBLrel. 23, Last annotation update)			
DE	Semaphorin 3A.			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;			
OC	Xenopodinae; Xenopus.			
OX	NCBI_TaxID=8355;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Head;			
RT	Tannahill D., Nielsen J., Regan A.G.;			
RT	"Xenopus Semaphorin 3A.";			
RL	Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AY030051; AAK38166.1; -			
DR	InterPro; IPR003599; Ig.			
DR	InterPro; IPR007110; Ig-like.			
DR	InterPro; IPR003006; Ig_MHC.			
DR	InterPro; IPR001005; MYB DNA binding.			
DR	InterPro; IPR003659; Plexin-like.			


```

Db 621 AspGluLeuValAspGluArgLysIleLysThrGluHisGlyLeuLeuLeuArgThr 640
QY 2114 CTACACAGAGGATTTCAGCAATACCTCTCCATGCGGTGGAAACATGGGTTCATACAA 2173
Db 641 LeuLysLysArgAspSerGlyIleTyrTyrCysAsnAlaValGluHisGlyPheMetGln 660
QY 2174 ACTCTCTTAAGGTAACTGACCTGAAGTCATGACAGACATTTGGAAGAACTTCTTCAT 2233
Db 661 ThrLeuLeuValThrLeuGluIleLeuAspThrGluHisLeuAspGluLeuLeuHis 680
QY 2234 AAAGATGATGAT- --GGAGATGGCTCTAAGACCAAGAAATCTCCAATAGCATGACACT 2290
Db 681 LysGluAspGluGlyAspSerHisLysHisLysGluProSerAsnSerMetSerPro 700
QY 2291 ACCGAGAAGTCTGTACAGACATCTCATGACAGCTCATCAACACCCCAATCTCAACACG 2350
Db 701 ThrGlnLysIleTyrTyrArgAspPheMetGlnLeuLeuLeuLeuLeuLeuLeuLeu 720
QY 2351 ATGGATGATGCTGTGACACAGTTTGGAAAGGACCCGAAACACACGTCGGCAAGGCCA 2410
Db 721 MetAspGluPheCysGluGlnValTyrLysArgAspArgLysGlnArgGlnLysAsn 740
QY 2411 GGACATACCCCGAGGAAACAGTAACAAATGGAGACACTTACAGAAATAAGAAAGGTAGA 2470
Db 741 GlyAsnValGlnValSerAsnThrLysIleLysHisLeuGlnGluAsnLysLysGlyArg 760
QY 2471 AACAGGAGACCCACCAATTTGAGAGGCGCACCCAGAGTGTC 2512
Db 761 AsnArgArgThrHisGluPheGluArgAlaProArgSerVal 774

RESULT 2
QBH34
ID QBH34 PRELIMINARY; PRT; 777 AA.
AC QBH34;
DT 01-MAR-2003 (TREMREL. 23, Created)
DT 01-MAR-2003 (TREMREL. 23, Last sequence update)
DT 01-MAR-2003 (TREMREL. 23, Last annotation update)
DE Semaphorin 3D precursor homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney, and Skin;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
  60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK028900; BAC26185.1; -.
DR EMBL; AK052671; BAC35092.1; -.
SQ SEQUENCE 777 AA; 89548 MW; 5450D8D45D1BDABF CRC64;

```

Alignment Scores:

```

Pred. No.: 4,346-183 Length: 777
Score: 2258.00 Matches: 428
Percent Similarity: 72.24% Conservative: 147
Best Local Similarity: 53.77% Mismatches: 190
Query Match: 46.12% Indels: 31
DB: 11 Gaps: 13

```

US-09-774-490-1 (1-2709) x QBH34 (1-777)

```

QY 148 ACNATACAGGAGGAGACTAAGACGAAAGGACCTACAGCTCTCCAGCATGGGCTG 207
Db 4 ThrLysAspGluAsnProArgSerArgSerGlnAspLeu---HisLeuPheHis-AlaTr 22
QY 208 GTTAACTAGGATTGTC---TGCTTTTCTGGGAGTATTACTTACAGCAAGCAAACTA 264
Db 22 pMetLeuLeuMetThrValLeuPheLeuProValThrGluThrSer----- 38

```

```

QY 265 TCAGAAATGGGAAGAACAAATGTGCCAAGCTGAAATTTATCTACAAAGAAATGTTGGAATC 324
Db 39 -----LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuSe 55
QY 325 CAACAATGTGATCATTTCATGCTTGGCCCAACAGCTCCAGTTATCATCTCTCTTTT 384
Db 55 rAsnThrCysIleProPheLeuGlySerSerGlyLeuAspPheGlnThrLeuLeuSe 75
QY 385 GGATGAGAAACCGAGTAGCTGTATGTGGAGCAAGCATCATATTTTTCATTCGACT 444
Db 75 uAspGluArgGlyLeuLeuLeuGlyAlaLysAspHisValPheLeuLeuSeLe 95
QY 445 GGTAAATATC---AAGATTTTCAAAAGATTGTGTGCCAGTATCTTACACAGAGAGA 501
Db 95 uValAspLeuAsnLysAsnPheLysIleTyrTyrProAlaAlaLysGluArgValG1 115
QY 502 TGAATGCAAGTGGCTGGAAGACATCTGAAAGATGTGTATTTTCATCAAGTACT 561
Db 115 uLeuCysLysLeuAlaGlyLysAspAlaAsnAlaGluCysAlaAsnPheIleArgVal 135
QY 562 TAAGGCATATATCAGACTCACTTGTACGCTGTGGAAACGGGGCTTTTCATCAATTTG 621
Db 135 uGlnProTyrAsnLysThrHisValTyrValCysGlyThrGlyAlaPheHisProLeuCy 155
QY 622 CACCTACATTAATTTGACATCATCTCGAGACAATATTTTAACTGGAGAACTACA 681
Db 155 sGlyTyrIleAspLeuGlyAlaAsnLysGluLysLeuLysLeuAspThrHisAs 175
QY 682 TTTTGAACCGCGTGGAGAGTCCATATGACCTTAAGCTGTGTGACAGCATCTCTTT 741
Db 175 nLeuGluSerGlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerVal 195
QY 742 AATAGATGGAGATATATCTCTGAACTGACAGCTGATTTTATGGGGCAGAGCTTGCTAT 801
Db 195 tThrAspGluHisLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPh 215
QY 802 CTTCGAACTCTCGG-----CACCAACCCCAATCAGACAGACAGCATGATTC 852
Db 215 eThrArgSerLeuGlyLeuMetGlnAspHisSerIleArgThrAspIleSerGluHi 235
QY 853 CAGTGGCTCAATGATCCAAAGTTTCATGAGCCACCTCATCTCAGAGAGTGACATCC 912
Db 235 sHisTrpLeuAsnGlyAlaLysPheIleGlyThrPheProIleProAspThrThrAsnPr 255
QY 913 TGAAGATCACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGAGAACTCTGG 972
Db 255 oAspAspLysIleTyrPhePheArgGluSerSerGlnGluGlySerThrSerAs 275
QY 973 AAAAGCTACTCACCTAGATAGATATGCAAGATGACCTTTGAGGGGACAGAG 1032
Db 275 pArgSerIleLeuSerArgValGlyArgValCysLysAsnAspValGlyGlnArgSe 295
QY 1033 TCTGTGATTAATGAGACATCTCCCAAGCTCGTCTGATTTGCTCAGTGGCCAGTCC 1092
Db 295 rLeuIleAsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerIleProGlySe 315
QY 1093 AAATGGCATTCACACTCATTTTGTAGTGAAGTGCAGAGTGTATTCTTAATGAATTTAAAGA 1152
Db 315 rAspGlyAlaAspThrHisPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAs 335
QY 1153 TCCTAAATAATCCAGTTGTATATGAGTGTTTTACAGCTTCCAGTAACATTTTCAAGGGATC 1212
Db 335 pGluArgAsnProValValTyrGlyValPheThrThrThrSerSerIlePheLysGlySe 355
QY 1213 AGCGTGTATGTATAGCATGATGTGAGAGGGTGTTCCTTGGTCCATATGCCCA 1272
Db 355 rAlaValCysValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHi 375
QY 1273 CAGGATGGAGCCCAACTATCAATGGGTGCCTTATCAAGGAAGAGTCCCTTATCCACGCCC 1332
Db 375 sLysGluSerAlaAspHisArgTrpValGlnTyrAspGlyArgIleProTyrProArgPr 395
QY 1333 AGGAACCTTGCCAGCAAAACATTTTGTGGT---TTTGACTCTACAAAGGACCTTCCTGA 1389

```

```
Db 395 oGlyThrCysProSerLysThrTyAspProLeuLeuLysSerThrArgAspPheProAs 415
QY 1390 TGATGTTATACCTTGGCAAGATCATCCAGCCATGATCAATCCAGTGTTCCTATGAA 1449
Db 415 pAspValIleSerPheIleArgHisProValMetTyrLysSerValTyrProValAl 435
QY 1450 CAATCGCCCAATAGTATCAAAACGATGTAATATCAATTTACAAATTTGCTGAGA 1509
Db 435 aGlyAlaProThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAs 455
QY 1510 CCGAGTGCATGCAGAGATGAGCAGTATGATGTTATGTTATCGGAACAGATGTTGGAC 1569
Db 455 phisValValAlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyTh 475
QY 1570 CTTCTTAAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATATTTAGAGAGTTCTGCT 1629
Db 475 rValLeuLysValValSerIleSerLysGluLysTrp--AsnMetGluGluValValle 494
QY 1630 GGAAGAAATGACAGTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTCCACTAA 1689
Db 494 uGluGluLeuGlnValPheLysHisProThrAlaIleLeuAsnMetGluLeuSerLeuLy 514
QY 1690 GCAGCAACAACTATATATTGGTTCAACGGCTGGGTTGCCAGCTCCCTTTACACGGTG 1749
Db 514 sGlnGlnGlnLeuTyrValGlySerTrpAspGlyLeuValGlnLeuSerLeuHisArgCy 534
QY 1750 TGATATTACGGGAAGCGTGTGAGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1809
Db 534 sAspThrTyrGlyLysAlaCysAlaAspCysLysLeuAlaArgAspProTyrCysAlaTr 554
QY 1810 GCATGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1869
Db 554 pAspGlyAsnAlaCysSerArgTyrAlaProThrSerLysArgAlaArgGlnAs 574
QY 1870 TATAAAGAAATGAGACCCACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1929
Db 574 pValLysTyrGlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis-- 593
QY 1930 CCACAGCCCTGAAGAGAGATCATCTATGCTGTAGAGATAGTAGACATATTTTGAATG 1989
Db 594 -GluThrAlaAspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCy 613
QY 1990 CAGTCCGAGTCCGAGAGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2049
Db 613 sIleProLysSerGlnGlnAlaSerValGluTrpTyrIleGlnArgSerGlyAspGluHi 633
QY 2050 AAAAGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2109
Db 633 sArgGluGluLeuLysProAspGluArgIleIleLysThrAspTyrGlyLeuLeuLeAr 653
QY 2110 TAGTCTAACAGAGATTCAGGCAATTAATCTCTGCCATGCGGTGGGAACATGGGTTGAT 2169
Db 653 gSerLeuGlnLysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheI 673
QY 2170 ACAACTCTTCTTAAGTAACCTCGAGTCAATTCAGACAGACATTTGGAGAACTTCT 2229
Db 673 ehistrIleValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsn---Th 692
QY 2230 TCATAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2289
Db 692 rGlnArgAlaGluTyrGlnGluGlyGlnValLysAspLeuLeuAlaGluSer----- 709
QY 2290 TAGCCAGAAGGTCTGGTACAGACTTCATGAGCTCATCAACACCCCAATCTCAACAC 2349
Db 710 -----ArgLeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---Se 726
QY 2350 GATGATGATGTTCTGTGAACAAGTTTGGAAAGGACCGAAACACGTCGGCAAGGCC 2409
Db 726 rIleuAspGlnTyrCysGluGlnMetTyrTyrLysGluLysArgGlnArgAsnLys-- 745
QY 2410 AGGACATACCCCGGGAACAGTAACAAATGGAAGCACTTTACAAGAAAAATAAGAAAGGTAG 2469
```

```
746 -----GlySerProLysTrpLysHisMetGlnGluMetLysLysLys 760
QY 2470 AAACAGGAGGACCCAC---GAATTTGAGAGCGCCACCCAGGAGTCTC 2512
Db 760 gasnArgArgHisArgAspLeuAspGluLeuGlnArgSerVal 775
RESULT 3
QBMP6
ID QBMP6 PRELIMINARY; PRT; 777 AA.
AC QBMP6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Senaphorin 3D precursor homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK031704; BAC27522.1; -.
SQ SEQUENCE 777 AA; 89562 MW; B7204D82288B89CE CRC64;
Alignment Scores:
Pred. No.: 1,48-182 Length: 777
Score: 2252.00 Matches: 427
Percent Similarity: 72.11% Conservative: 147
Best Local Similarity: 53.64% Mismatches: 191
Query Match: 46.00% Indels: 31
DB: 11 Gaps: 13
US-09-774-490-1 (1-2709) x QBMP6 (1-777)
QY 148 ACAATACAGGAGGAAGACTAAAGCAAGGACCTTACAGCGCTGTCAGCATGGGCTG 207
Db 4 ThrLysAspGluAsnProArgSerArgSerGlnAspLeu---HisLeuPheHis-AlaTr 22
QY 208 GTTAACTAGGATTCTC---TGTCTTTCTGGGAGTATTACTTACAGCAAGCAAACTA 264
Db 22 pMetMetLeuIleMetThrValLeuPheLeuProValThrGluThrSer----- 38
QY 265 TCAGAAATGGGAAGAACCAATGTCCCAAGCTGAAATTTATCTTACAAAGAAATTTGGAATC 324
Db 39 -----LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuLe 55
QY 325 CAACAATGTGATCACTTTCAATGCTTGGCAACAGCTCCAGTTATCATACCTCTTTT 384
Db 55 rAsnThrCysIleProPheLeuGlySerSerGlyLeuAspPheGlnThrLeuLe 75
QY 385 GGATGAGAACCGAGTAGGCTGTATGTTGGCAAGAGATCATATTTTTCATTCGACT 444
Db 75 uAspGluGluArgGlyLeuLeuLeuGlyAlaLysAspHisValPheLeuLeuSerLe 95
QY 445 GTTTAAATATC---AAGGATTTTCAAAAGATTTGTGCGCAGTATCTTACACCAAGAGAGA 501
Db 95 uValAspLeuAsnLysAsnPheLysLysIleTyrTrpProAlaLysGluArgValG 115
QY 502 TGAATGCAAGTGGCTGGAAAGACATCCTCTGAAAGATGCTTAATTTTCAAGGTACT 561
Db 115 uLeuCysLysLeuAlaGlyLysAspAlaAsnAlaGluCysAlaAsnPheIleArgVal 135
QY 562 TAAGGCATATATCAGACTCATCTTGTACGCTGTGGCAACCGGGCTTTTCAATCAATTG 621
Db 135 uGlnProTyrAsnLysThrHisValTyrValCysGlyThrGlyAlaPheHisPLeuLeu 155
QY 622 CACCTACATTGAAATTTGGACATCATCTCGAGGACAATATTTTAAAGCTGGAGAACTACA 681
```

Db 155 sGlyTyrIleAspLeuGlyAlaAsnLysGluGlnLeuIlePheLysLeuAspThrHisAs 175
QY 682 TTTTCAAAACGGCCGTGGAGAGTCCATATACACCTAGCTGCTCAGACAGCTCCCTTTT 741
Db 175 nLeuGluSerGlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMe 195
QY 742 AATAGATGAGAAATATATCTCGAACTGACAGTCAATTTATGCGGCGAGACTTTGCTAT 801
Db 195 tThrAspGluHisLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPh 215
QY 802 CTTCCGAACCTCTGGG-----CACCACCAACCAATCAGGACAGACGACGATGATC 852
Db 215 eThrArgSerLeuGlyLeuMetGlnAspHisSerIleArgThrAspIleSerGluHi 235
QY 853 CAGGTGGCTCAATGATCCAAAGTTCATTAGTCCACCTCATCTCAGAGAGTGACAATCC 912
Db 235 shiStrLeuAsnGlyAlaLysPheIleGlyThrPheProIleProAspThrTyrAsnPr 255
QY 913 TGAAGATGACAAAGTATATCTTTCTTCGTGAAATGCAATAGATGAGAACACTCTGG 972
Db 255 oAspAspAspLysIleTyrPhePhePheArgGluSerSerGlnGluGlySerThrSerAs 275
QY 973 AAAAGCTACTCAGCTAGATAGTTCAGATATGCAAGATGACTTTGGAGGCGACAGAG 1032
Db 275 pArgSerIleLeuSerArgValGlyArgValCysLysAsnAspValGlyGlnArgSe 295
QY 1033 TCTGGTGAATAAATGACAACTTCTCAAAGCTCGTGTGATTTCTCAGTGCCAGGTCC 1092
Db 295 rLeuIleAsnLysTyrThrPheLeuLysAlaArgLeuIleCysSerIleProGlySe 315
QY 1093 AAATGGCATTGACATCTATTTTGATGAACTGAGGATGATTCCTTAATGAATTTAAAGA 1152
Db 315 rAspGlyAlaAspThrHisPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAs 335
QY 1153 TCCTAAATCCAGTGTATATAGTGTGTGAGAGGGTGTCTTGTGTCATATGCCCA 1272
Db 335 pGluArgLysProValValTyrGlyValPheThrThrThrSerSerIlePheLysGlySe 355
QY 1213 AGCCGTGTGTATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1272
Db 355 rAlaValCysValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHi 375
QY 1273 CAGGATGAGACCAACTATCAATGAGTGTGATGATGATGATGATGATGATGATGATGAT 1332
Db 375 sLysGluSerAlaAspHisArgTyrValGlnTyrAspGlyArgIleProTyrProArgPr 395
QY 1333 AGGAATCTTCCAGCAAAACATTTGGTGGT---TTTGACTCTCAAAAGGACCTTCCTGA 1389
Db 395 oGlyThrCysProSerLysThrTyrAspProLeuIleLysSerThrArgAspPheProAs 415
QY 1390 TGATGTTTAACTTTGCAAGAGTCAATCCAGCCATGTACATCCAGTGTTCCTATGAA 1449
Db 415 pAspValIleSerPheIleArgArgHisProValMetTyrLysSerValTyrProValAl 435
QY 1450 CAATCGCCCAATAGTATCAAAACCGAGTGAATATCAATTTACAAATTTGCTGAGA 1509
Db 435 aclyAlaProThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAs 455
QY 1510 CCGAGTGTATGAGAGATGAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1569
Db 455 pHisValValAlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyTh 475
QY 1570 CTTCTTAAAGTAGTTTCAATTTCCTAAGGAGACTTGGTATGATTTAGAGAGGTTCTGCT 1629
Db 475 rValLeuLysValValSerIleSerLysGluLysTyr---AsnMetGluGluValValLe 494
QY 1630 GGAAGAAATGACAGTTTTCGGGAACCGACTGCTATTTTTCAGCAATGGAGCTTTTCCACTAA 1689
Db 494 uGluGluLeuGlnValPheLysHisProThrAlaIleLeuAsnMetGluLeuSerLeuLy 514
QY 1690 GCAGCAACACTATATATTTGTTCAACGGCTGGGTGCTCCAGCTCCCTTTTACACCGGTG 1749

Db 514 sGlnGlnGlnLeuTyrValGlySerTyrAspGlyLeuValGlnLeuSerLeuHisArgCy 534
QY 1750 TGATATTTTACCGGAAAGCGTGTGCTGAGTGTGCTGCTCGCCGAGACCTTACTGTGCTTG 1809
Db 534 sAspThrTyrGlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTr 554
QY 1810 GGATGGTCTCATGTTCTCGCTATTTTCCACTGCAAGAGAGCGCACAGACGACAGA 1869
Db 554 pAspGlyAsnAlaCysSerArgTyrAlaProThrSerLysArgAlaArgArgGlnAs 574
QY 1870 TATAAGAAATCGAGACCCACTGACTCTGTTGTAGACTTACCATGATATATCACCATTGG 1929
Db 574 pValLysTyrGlyAspProIleThrGlnCysTyrAspIleGluAspSerIleSerHis-- 593
QY 1930 CCACAGCCCTCAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGACATTTTGGAAATG 1989
Db 594 -GluThrAlaAspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCy 613
QY 1990 CAGTCCGAGTCCGAGAGAGCGCTGCTTATTTGGAATTTCCAGAGCGGCAATGAAGAGG 2049
Db 613 sIleProLysSerGlnGlnAlaSerValGluTyrTyrIleGlnArgSerGlyAspGluHi 633
QY 2050 AAAAGAGAGATCAGATGAGTATCATATCATCAGGACAGATCAAGGCTTCTGCTAGC 2109
Db 633 sArgGluLeuLeuLysProAspGluArgIleIleLysThrAspTyrGlyLeuLeuIleAr 653
QY 2110 TAGTCTCAACAGAGATTCAGGCAATTCCTCTGCCATCGGTGGAACTGGTTCAT 2169
Db 653 sSerLeuGlnLysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIl 673
QY 2170 ACAACTCTTCTTAAGTAACTGAGTACCTGGAAGTCAATTCAGACAGAGCATTTGGAAACTTCT 2229
Db 673 eHisThrIleValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsn---Th 692
QY 2230 TCATAAGATCATGATGAGATGCTCTAAGACCAAGAAATGTCCATATGATCATCACACC 2289
Db 692 rGlnArgAlaGluTyrGlnGluGlyGlnValLysAspLeuAlaGluSer----- 709
QY 2290 TAGCCAGAAGCTGCTGACAGAGACTTCATCGAGTTCATCAACACCCCAATCTCAACAC 2349
Db 710 -----ArgLeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---Se 726
QY 2350 GATGATGAGTTCGTGAAACAAAGTTTGAAGAGGCGGAAACCAACCTCGGCAAGGCC 2409
Db 726 rLeuAspGlnTyrCysGluGlnMetTyrTyrLysLysArgGlnArgAsnLys-- 745
QY 2410 AGGACATACCCCGGAGACAGTAACTAAGTGAAGTCTTACAGAAATAAGAAAGGTAG 2469
Db 746 -----GlySerProLysTyrLysHisMetGlnGluMetLysLysLysLeAr 760
QY 2470 AACAGGAGGACCCAC---GAATTTGAGAGGCGCCACCCAGGAGTGTCT 2512
Db 760 gAsnArgArgHisArgAspLeuAspGluLeuGlnArgSerVal 775

RESULT 4
Q8TB71 PRELIMINARY; PRT; 748 AA.
ID AC Q8TB71;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024220; AAH24220.1; -;
DR InterPro; IPR003599; Ig.

QY	2063	AGAGTGGATGATCATATCATCAGACAGATCAAGGCCTTCTGCTACGCTAGTCTACAAACAG	2122
Db	616	LeuAlaGluGluArgThrGluArgThrAlaArgGlyLeuLeuLeuArgLeuArgArg	635
QY	2123	AAGGATTACAGGCAATTACCTCTGCATCGCGTGGAAACATGGTTCATACAAACTCTCTTT	2182
Db	636	ArgAspSerGlyValTyLeuCysAlaAlaValGluGlnGlyPheThrGlnProLeuArg	655
QY	2183	AAGGTAACCCCTGGAGTCATTGACACAGAGCATTTTGAAGAACCTCTTCATAAAGATGAT	2242
Db	656	ArgLeuSerLeuHisValLeuSerAlaThrGlnAlaGluArgLeu	670
QY	2243	GATGGAGATGGCTCTAGACCAAGAAGATGTCATAGCATGACACCTACCCAGAAGTTC	2302
Db	671	-----AlaArgAlaGluGluAlaProAlaAlaProGlyProLysLeu	686
QY	2303	TGGTACAGAGACTTCATCGAGCTCATCAACCCCAATCTCAACGATGGATGAGTTC	2362
Db	687	TrpTyArgAspPheLeuGlnLeuVal	695
QY	2363	TGTGAACAAGTTTGGAAAAGGGACCGAAAAACAACGTCGGCAAGGCCAGGACATACCCCA	2422
Db	696	-----GluProGly---GlyGly	700
QY	2423	GGGNAACAGTACAAATCGAAG	2449
Db	701	GlySerAlaAsnSerLeuArgMetCysArgProGlnProAlaLeuGlnSerLeuProLeu	720
QY	2450	CAAGAAATAAGAAAGGTAGAAACAGGACGAGACCCAC	2497
Db	721	GluSerArgArgLysGlyArgAsnArgArgThrHisAlaProGluProArgAlaGluArg	740
QY	2498	GCACCCAGGAGT 2509	
Db	741	GlyProArgSer 744	
RESULT 5			
Q8TDV7			
ID	Q8TDV7	PRELIMINARY;	PRT; 754 AA.
AC	Q8TDV7		
DT	01-JUN-2002	(TrEMBLrel. 21, Created)	
DT	01-JUN-2002	(TrEMBLrel. 21, Last sequence update)	
DT	01-MAR-2003	(TrEMBLrel. 23, Last annotation update)	
DE	Semaphorin 3B.		
GN	SEMA 3B.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RA	SEQUENCE FROM N.A.		
RA	Koyama N.;		
RT	"semaphorin 3B (SEMA3B) cDNA.";		
RL	Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AB083186; BAB88870.1; -		
DR	InterPro; IPR003599; Ig.		
DR	InterPro; IPR007110; Ig-like.		
DR	InterPro; IPR003006; Ig_MHC.		
DR	InterPro; IPR003659; Flexin-like.		
DR	InterPro; IPR001627; Sema.		
DR	Pfam; PF00047; Ig; 1.		
DR	Pfam; PF01403; Sema; 1.		
DR	SMART; SM00409; IG; 1.		
DR	SMART; SM00423; PSI; 1.		
DR	SMART; SM00630; Sema; 1.		
DR	PROSITE; PSS0835; IG_LIKE; 1.		
QY	SEQUENCE 754 AA; 83691 MW; E61FD08C04B9A68E CRC64;		
Alignment Scores:			
Pred. No.:	7.21e-170		Length: 754
Score:	2102.50		Matches: 403
Percent Similarity:	66.79%		Conservative: 124

Best Local Similarity:	51.08%	Mismatches:	185
Query Match:	42.94%	Indels:	77
DB:	4	Gaps:	13
US-09-774-490-1 (1-2709) x Q8TDV7 (1-754)			
QY	221	GTCTGCTTTTCTGGGAGTATTACTTACACCAAGACAAACTATCAGAAATGGGAGAAC	280
DB	13	LeuAlaLeuLeuTrpAlaValGlyLeuGlySerAlaAla-----Pro	26
QY	281	AATGTGCCAGCGTGAATATCTCTACAAAGAAATGTTGGAATCCAACAATGTGTATCAT	340
DB	27	SerProProArgLeuArgLeuSerPheGlnGluLeuGlnAlaTrpHisGlyLeuGlnThr	46
QY	341	TTCAATGGCTTGGCCAAACAGCTCCAGTTTATCATCTTCTCTTTGGATGAGAAACGGAGT	400
DB	47	PheSer---LeuGluAtrGthrCysCysTyrGlnAlaLeuLeuValAspGluGluArgGly	65
QY	401	AGGCTGTATGTTGGAGCAAGGATCATATTTTCATTCGACCTGGTTAAATATC---AAG	457
DB	66	ArgLeuPheValGlyAlaGluAsnHisValAlaSerLeuAsnLeuAspAsnIleSerLys	85
QY	458	GATTTTCAAAGATTGTGSCCAGTATCTTACACCAGACAGATGAATGCAAGTGGCGCT	517
DB	86	ArgAlaLysLysLeuAlaTrpProAlaProValGluTrpArgGluGluCysAsnTrpAla	105
QY	518	GGAAAAGACATCTCTGAAAGAAATGTCTAAATTCATCAAGGTACTTAAAGGCATATAATCAG	577
DB	106	GlyLysAspIleGlyThrGluCysMetAsnPheValLysLeuLeuHisAlaTyrAsnArg	125
QY	578	ACTCATTTGACGCTGTGGAAACGGGGCTTTTCATCCAAATTTGCACTCATGTAATAAT	637
DB	126	ThrHisLeuLeuAlaCysGlyThrGlyAlaPheHisProThrCysAlaPheValGluVal	145
QY	638	GGACATCATCTGAGGCAATATTTTAAAGCTGTGAGAACTCACATTTTGAAACGGCGGT	697
DB	146	GlyHisArgAlaGluGluProValLeuArgLeuAspProGlyArgIleGluAspGlyLys	165
QY	698	GGGAAGAGTTCATATGACCCCTAAAGCTGTGACAGCATCCCTTTTAAATAGATGGAGAA	757
DB	166	GlyLysSerProTyrAspProArgHisArgAlaAlaSerValLeuValGlyGluLeu	185
QY	758	TACTCTGGAACTGCAGCTGATTTTATGGGGCGAGACTTTGCTATCTTCCCGAAGCTTGGG	817
DB	186	TyrSerGlyValAlaAlaAspLeuMetGlyArgAspPheThrIlePheArgSerLeuGly	205
QY	818	CACCACCACCAATCAGGACAGACGATGATTCAGGTGGCTCAATGATCCAAAGTTC	877
DB	206	GlnArgProSerLeuArgThrGluProHisAspSerArgTrpLeuAsnGluProLysPhe	225
QY	878	ATTAGTGCCCACTCATCTCAGAGAGTGACAATCCTGAAGATGACAAAGTATCTTTTTC	937
DB	226	ValLysValPheTrpIleProGluSerGluAsnProAspAspLysIleTyrPhePhe	245
QY	938	TTCCGTGAAATGCAATAGATGGAGAACACTCT---GGAAAAGTACTACGCTAGATA	994
DB	246	PheArgGluThrAlaValGluAlaAlaProAlaLeuGlyArgLeuSerValSerArgVal	265
QY	995	GGTCAGATATGCAAGATGACTTTGGAGGGGACAGAACTCTGTGTGAATTAATGACACA	1054
DB	266	GlyGlnIleCysArgAsnAspValGlyGlyLysArgSerLeuValAsnLysTrpThr	285
QY	1055	TTCCCTCAAAGCTGTGATTTGCTCAGTCGCCAGGTCCAAATGCATTTGACACATCTTT	1114
DB	286	PheLeuLysAlaArgLeuValCysSerValProGlyValGluGly---AspThrHisPhe	304
QY	1115	GATGAACTG-----CAGATGTATTTCTTAATGAACTTTAAAGATCCCTAAA	1159
DB	305	AspGlnLeuArgProPheProAlaGluAspValPheLeuLeuSerSerArgAspHisArg	324
QY	1160	AATCCAGTTGTATGGAGTGTTTACGACTTCCAGTAACTTTTCAAGGATCAGCCGTG	1219
DB	325	ThrProLeuLeuTrpAlaPheSerThrSerSerIlePheGlnGlySerAlaVal	344

Alignment Scores:	
Pred. No.:	7.21e-170
Score:	2102.50
Percent Similarity:	66.79%
Length:	754
Matches:	403
Conservative:	124

Db 759 LeuArgSerLysAlaGluHisPheArgLeuProArg 770

RESULT 7

Q8QUG9 PRELIMINARY; PRT; 756 AA.

AC Q8QUG9; 01-JUN-2002 (TREMBlrel. 21, Created)

DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)

DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)

DE Semaphorin 3F.

GN SEMA3F.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OC NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RA Watanabe Y., Nakamura H.;

RT "Axon guidance of the trochlear nerve by Sema 3F along mid-hindbrain boundary.";

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB072930; BAB8691.1; -

DR InterPro; IPR007110; IG-like.

DR InterPro; IPR003598; IG_c2.

DR InterPro; IPR003006; IG_MHC.

DR InterPro; IPR003659; Plexin-like.

DR InterPro; IPR001627; Sema.

DR Pfam; PF00047; ig; 1.

DR Pfam; PF01403; Sema; 1.

DR SMART; SM00408; IGC2; 1.

DR SMART; SM00423; PSI; 1.

DR SMART; SM00630; Sema; 1.

DR PROSITE; PS00835; IG-LIKE; 1.

KW Immunoglobulin domain.

SQ SEQUENCE 756 AA; 85196 MW; C36754C02541ED88 CRC64;

Alignment Scores:

Pred. No.: 128e-161 Length: 756

Score: 2005.50 Matches: 380

Percent Similarity: 68.16% Conservative: 138

Best Local Similarity: 50.00% Mismatches: 217

Query Match: 40.96% Indels: 25

DB: 13 Gaps: 11

US-09-774-490-1 (1-2709) x Q8QUG9 (1-756)

QY 227 CTTTCTGGGAGTATTACTACAGCAGCAAACTATCAGAATGGGAAGAACAAATGTG 286

Db 7 LeuLeuTriaAlaThrLeuLeuThrLeuGlyTyrArgAlaAlaHisGlyLysAspGlyVal 26

QY 287 -----CDAAGCTGAATATCTCTACAAAGAAATGTTGGATCCCAATGTGATC 337

Db ProProThrProArgValGlnLeuSerPheLysGluLeuLysAlaThrGlyThrAlaHis 46

QY 338 ACTTCAATGGCTGGCCCAAGCTCCAGTATATACCTCTTGGATCAGGAACGG 397

Db PhePheAsnPheLeuLeuAsnSerAspTyrArgIleLeuLeuLysAspGluAspHis 66

QY 398 AGTAGGCTGTATGTTGGCAAGGATCACATATTTTCATTCGACTGCTTAATATCAAG 457

Db AspArgMetTyrValGlySerLysAspTyrValLeuSerLeuAspLeuHisAspIleAsn 86

QY 458 GATTTTCAAAAGATTGTG---TGGCCAGTATCTTACACCAAGAGAGATGAATGCAAGTGG 514

Db ArgGluProLeuIleIleHisTyrProAlaSerGlnGlnArgIleGluGluCysIleLeu 106

QY 515 GCTGGAAGAAGATCTCTGAAAGAGTGTCTAATTTTCATCAAGTACTTAAGGCATATAAT 574

Db SerGlyLysAsnSerAsnGlyGluCysGlyAsnPheIleArgLeuIleGlnProtrpAsn 126

QY 575 CAGACTCATTGTACGCTGTGGACGGGGCTTTTTCATCCCAATTTGCACCTACATTGAA 634

Db 127 ArgThrHisLeuTyrValCysGlyThrGlyAlaTyrAsnProIleCysAlaPheIleAsn 146

QY 635 ATTGGACATCATCTGAGGACAATATTTTAAAGCTGGAGAACTCACATTTTAAAACCGC 694

Db 147 ArgGlyArgLysAlaGlnAspTyrIlePheTyrLeuGluProAspLysLeuGluSerGly 166

QY 695 CGTGGGAAGAGTCCATATGACCTTAAGCTGTGACAGCATCCCTTTTAAATAGATGAGAA 754

Db 167 LysGlyLysCysSerTyrAspProLysValAspThrValSerAlaLeuIleAsnGluGlu 186

QY 755 TTATCTCTGGAAGTCTGAGTATTTATGGGGGAGAGCTTTTGTCTATCTTCGAACTCTT 814

Db 187 LeuTyrAlaGlyValTyrIleAspPheMetGlyThrAspAlaAlaIlePheArgThrMet 206

QY 815 GGGCACCACCAATCAGCAGCAGCATGATTCAGGTGGCTCAATGATGATCAAG 874

Db 207 GlyLysGlnThrAlaMetArgThrAspGlnTyrAsnSerArgTyrLeuAsnAspProala 226

QY 875 TTCATTAGTCCACCTCATCTCAGAGAGTCACCAATCTCGAAGATGACAAAGTATATCTT 934

Db 227 PheValArgAlaGlnLeuIleProAspSerSerGluArgAsnAspAspLysLeuTyrPhe 246

QY 935 TTCTTCGCTGAAATGCAATAGATGGAGAACACCTCTCGAAAGAGTACTCACGCTAGATA 994

Db 247 PhePheArgGluLysSerAlaAspAlaProLeuSer---ProGlyValTyrSerArgIle 265

QY 995 GGTGATATGCAAGATGACTTTGGGGGACAGAGTCTGGTGAATAAATGGAACA 1054

Db 266 GlyArgIleCysLeuAsnAspAspGlyGlyHisCysCysLeuValAsnLysTyrPheThr 285

QY 1055 TTCTCAAAAGCTCGTCTGATTTGCTCAGTGGCCAGTCCAAATGGCATTGACACTCATTT 1114

Db 286 PheLeuLysAlaArgLeuValCysSerValProGlyProAspGlyIleGluThrHisPhe 305

QY 1115 GATCACTGCAAGATGATCTCTAATGAACCTTTAAAGATCTCTAAATTCAGTGTATAT 1174

Db 306 AspGluLeuGlnAspValPheIleGlnThrGlnAspThrLysAsnProValIleTyr 325

QY 1175 GGAGTGTTCAGCTTCCAGTAAACATTTTCAAGGATCAGCGGTGTGTATGATATGATG 1234

Db 326 AlaValPheSerAlaSerGlySerValPheLysGlySerAlaValCysValTyrSerMet 345

QY 1235 AGTGTATGCAAGAGGTGTTCTTCTGCTCATATGCCACAGGATGGAACCAACTATCAA 1294

Db 346 AlaAspIleArgMetValPheAsnArgProPheAlaHisLysGluGlyProAsnTyrGln 365

QY 1295 TGGTGTCTTATCAAGGAAGAGTCCCTTATCCAGGCCAGGAACCTTGTCCAGCAACA 1354

Db 366 TrpMetProTyrThrGlyLysMetProTyrProArgProGlyThrCysProGlyGlyThr 385

QY 1355 TTT---GGTGGTTTGTACTCTACAAAGGACCTTCTGATGATGTTATTAACCTTTCGAAGA 1411

Db 386 PheThrProSerMetLysSerThrLysAspTyrProAspGluValIleAsnPheMetArg 405

QY 1412 AGTCATCCAGCATGTACAATCCAGTGTTCCTATGATCAATCGCCCAATAGTGTATCAA 1471

Db 406 SerHisProLeuMetTyrHisAlaValTyrProAlaHisArgGlnProLeuValArg 425

QY 1472 ACGGATGTAATATCAATTTACAAATTCGTCGAGCCAGGATGCGAGAGATGGA 1531

Db 426 ThrAsnValAsnTyrArgPheThrThrIleAlaValAspGlnValAspAlaAlaAspGly 445

QY 1532 CAGTATGATGTATGTTTATCGAACAGATCTTGGACCGTCTTAAAGTAGTTTCAATT 1591

Db 446 ArgTyrGluValLeuPheLeuGlyThrAspArgGlyThrValGlnLysValIleValLeu 465

QY 1592 CTAAGGAGACTTGTGTATGATTTAGACAGAGTCTTCTGGAAGAAATGACAGTTTTCGG 1651

Db 466 ProArgAspAsp---MetGluThrGluGluLeuMetLeuGluGluValPheLys 484

QY 1652 GAACCGACTGTATTTCCAGCAATGAGCTTTCCTAAGCAGCAACAATATATATGTT 1711

Db 485 ValProAlaProIleLysMetMetThrIleSerLysArgGlnGlnLeuTyrValSer 504

```
QY 1712 TCAACGGCTGGGTTGCCAGCTCCCTTTACACCGGTGTGATATTATTACGGGAACGGTGT 1771
Db |||||
505 SerAlaValGlyValThrHisLeuAlaLeuHisArgCysAspValTyrGlyGluAlaCys 524
QY 1772 GCTGAGTGTGCTCGCCGAGACCTTACTGTCTGGTGGTGGTGTCTGCATGTTCTGCG 1831
Db |||||
525 AlaAspCysLeuAlaArgAspProTyrCysAlaTyrAspGlyLysAlaCysSerArg 544
QY 1832 TATTTTCCCACTGCAAGAGACGCAACAGACGACCAAGATATAGAAATGGAGACCACTG 1891
Db |||||
545 TyrSerAlaSerSerLysArgSerArgArgGlnAspValArgHisGlyAsnProMet 564
QY 1892 ACTCTACTGTTACAGATTACCATGATATACCATGCGCCACGACCCCTGAAAGAGAAATC 1951
Db |||||
565 ArgGlnCysArgGly-----TyrAsnSerAsnAlaAsnLysAsnThrValGluAlaVal 582
QY 1952 ATCTATGTTAGAGAAATAGTAGACATTTTGGATGCGATCCGAAAGTCCGAGAGCGG 2011
Db |||||
583 GlnTyrGlyValGluGlySerThrAlaPheLeuGluCysGlnProArgSerProGlnAla 602
QY 2012 CTGTGCTATTGGCAATTCAGAGGGCAATTAAGAGCGCAAGAAAGAGATCAGAGTGGAT 2071
Db |||||
603 ThrValLysTrpLeuLeuGlnLysAspAsnSerArgArgGlyGluLeuArgThrGlu 622
QY 2072 ---GATCATATCATCAGACAGATCAAGCCCTTCTGCTACGTAGTCTTACAACAGAGGAT 2128
Db |||||
623 GlyGlyArgAlaLeuArgThrGluGlnGlyLeuLeuLeuArgAlaLeuGlnLeuSerAsp 642
QY 2129 TCAGGCAATACCTCTGCATGCGGTGGAACATGGGTTTCATACAACTCTTCTTAAGGTA 2188
Db |||||
643 SerGlyLeuTyrSerCysThrAlaThrGluAsnAsnPheLysHisThrValThrLysVal 662
QY 2189 ACCCTGGAGTCAATGACACAGAGCATTTGGAAGAACTTCTTCAATAAGATGATGATGA 2248
Db |||||
663 GlnLeuArgValLeuAlaAlaArgAlaValHisAlaValLeuGlnGlyGlu----- 680
QY 2249 GATGCTCTAGAACCAAGAAATGTCATAGCATGACACCTPAGCCAGAGGTCTGTGTA 2308
Db |||||
681 -----LeuProProAlaAlaLeuProGlyAlaProThrProArg-----Tyr 694
QY 2309 AGAGACTTCATGCTCATCAACCCCACTTCAACACGATGATGATGATGATGATGATGAT 2368
Db |||||
695 GlnAspLeuLeuGlnLeuThrArgProGluLeuGlyLeuLeuAspGlnTyrCysGln 714
QY 2369 CAAGTTTGGAAAAGGACCGGAAACACAGCTCGGCAAGGCCAGGA---CATACCCCGGG 2425
Db |||||
715 GlyPheTrp-----ArgProProAlaProGlyProProGluProLeu 728
QY 2426 AACAGTAAACAAATGGAACACTTACAGAAATAGAAAGGTAGAAACAGGAGGACCCAC 2485
Db |||||
729 AlaAlaLeuLysAlaLysGluLeuGlnAspGlnLysLysProArgSerArgArgAsnHis 748

RESULT 8
Q96GX0 PRELIMINARY; PRT; 635 AA.
AC 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; BC009113; AA09113.1; --
DR InterPro; IPR003599; IG.
DR 240 PropheAlaHisLysGluGlyProMetHisGlnTrpValSerTyrGlnGlyArgValPro 259
```

```
DR InterPro; IPR003006; IG MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR001627; Sema.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS00835; IG LIKE; 1.
KW Hypothetical protein.
SQ SEQUENCE 635 AA; 70694 MW; D994099B476B9210 CRC64;

Alignment Scores:
Pred. No.: 3.9e-153 Length: 635
Score: 1905.50 Matches: 359
Percent Similarity: 68.30% Conservative: 102
Best Local Similarity: 53.19% Mismatches: 149
Query Match: 38.92% Indels: 65
DB: 4 Gaps: 10

US-09-774-490-1 (1-2709) x Q96GX0 (1-635)
QY 545 AATTTCATCAGGTACTTAAGGCATATTAATCAGACTCACTTGTAGCGCTGTGGACGGG 604
Db |||||
2 AsnPheValLysLeuLeuHisAlaTyrAsnArgThrHisLeuLeuAlaCysGlyThrGly 21
QY 605 GCTTTTCATCAATTTGCACCTACATTCGAAATTCGACATCATCTCGAGGCAATATTTT 664
Db |||||
22 AlaPheHisProThrCysAlaPheValGluValGlyHisArgAlaGluGluProValLeu 41
QY 665 AAGCTGGAGAACTCACATTTTGAACCGCGCTGGAGAGTCCATATAGACCCCTAAGCTG 724
Db |||||
42 ArgLeuAspProGlyArgGluGluAspGlyLysGlyLysSerProTyrAspProArgHis 61
QY 725 CTGACAGCATCCCTTTAATAGATGGAGATTAATCTCTGGAATCGAGCTGATTTATG 784
Db |||||
62 ArgAlaAlaSerValLeuValGlyGluLeuTyrSerGlyValAlaAlaAspLeuMet 81
QY 785 GGGCGAGACTTGTCTATCTCCGAACCTTTGGGCACACACCACTCAGACAGAGCAG 844
Db |||||
82 GlyArgAspPheThrIlePheArgSerLeuGlyGlnArgProSerLeuArgThrGluPro 101
QY 845 CATGATTCAGGTGGCTCAATGATCCAAAGTTCATTAGTCCACCTCATCTCAGAGAGT 904
Db |||||
102 HisAspSerArgTrpLeuAsnGluProLysPheValLysValPheTrpIleProGluSer 121
QY 905 GACATCTCTGAGATGACAAAGTATCTTTTCTCGTGAATGCAATAGATGGAGAA 964
Db |||||
122 GluAsnProAspAspLysIleTyrPhePheArgGluThrAlaValGluAlaAla 141
QY 965 CACTCT---GGAAAAGCTACTCAGCTAGATAGTGCAGATATCGAAGATGACTTTGGA 1021
Db |||||
142 ProAlaLeuGlyArgLeuSerValSerArgValGlnIleCysArgAsnAspValGly 161
QY 1022 GGGCACAGAGTCTGGTGAATAAATGGACAAACATCTCTCAAAGCTCGTCTGATTGTCTCA 1081
Db |||||
162 GlyGlnArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeuValCysSer 181
QY 1082 GTGCCAGGTCCAATGCGATTGACACTTTTGCATGAATGAGAGGATGATTTCTTATG 1141
Db |||||
182 ValProGlyValGluGly---AspThrHisPheAspGlnLeuGlnAspValPheLeuLeu 200
QY 1142 AACTTTTAAGATCCATAAATCCAGTGTATATGAGTGTGTTTACGACTTCCAGTAAACATT 1201
Db |||||
201 SerSerArgAspHisArgThrProLeuLeuTyrAlaValPheSerThrSerSer---Ile 219
QY 1202 TTCAAGGGATCAGCCGTGTGTATGATAGCATGATGTGAGAGGGTGTCTCTTGGT 1261
Db |||||
220 PheGlnGlySerAlaValCysValTyrSerMetAsnAspValArgAlaPheLeuGly 239
QY 1262 CCATATCCCAAGGATGAGCCCACTATCAATGGGTGCTTATCAAGAGAGTCCCC 1321
Db |||||
240 PropheAlaHisLysGluGlyProMetHisGlnTrpValSerTyrGlnGlyArgValPro 259
```

QY 1322 TATCAGCGCCAGGACTTGTCCCGACAAAACATTTGGTGGTTTTCACCTCTACAGGAC 1381
Db 260 TyrProArgProGlyMetCysProSerLysThrPheGlyThrPheSerSerThrLysAsp 279
QY 1382 CTTCTGATGATGTTTATACCTTTTGAAGAAGTCATCCAGCCATGTACAAATCCAGTGT 1441
Db 280 PheProAspValIleGlnPheAlaArgAsnHisProLeuMetTyrAsnSerValLeu 299
QY 1442 CTTATGAACAATCGCCCAATAGTATCAAAACGGATGTAAATTAATCAATTTACAAAT 1501
Db 300 ProThrGlyArgProLeuPheLeuGlnValGlyAlaAsnTyrPheThrGlnIle 319
QY 1502 GTCTGAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1561
Db 320 AlaAlaAspArgValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 339
QY 1562 GTTGGACCGCTTTCTAAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTAGAGAG 1621
Db 340 ValGlyThrValLeuLysValIleSerValProLysGlySerArgProSerAlaGluGly 359
QY 1622 GTTCTGCTGAAGAAATGACATGTTTTCGGGAACCGACTGCTATTTCAGCAATGAGCTT 1681
Db 360 LeuLeuLeuGluGluLeuHisValPheGluAspSerAlaAlaValThrSerMetGlnIle 379
QY 1682 TCCACTAAGCAGCAACACTATATATTGTTCAACGGCTGGGTTGCCAGCTCCCTTTA 1741
Db 380 SerSerLysArgHisGlnLeuTyrValAlaSerArgSerAlaValAlaGlnIleAlaLeu 399
QY 1742 CACCGTGTGATATTACCGGAAAGCGTGTGCTGAGTGTGCTCGCCGAGACCTTAC 1801
Db 400 HisArgCysAlaAlaHisGlyArgValCysThrGluCysCysLeuAlaArgAspProTyr 419
QY 1802 TGTGCTGGGATGTTCTGCATGTTCTGCTATTTTCCACATGCGAAGAGACGCAAGA 1861
Db 420 CysAlaTrpAspGlyValAlaCysThrArgPheGlnProSerAlaLysArgPheArg 439
QY 1862 CGACAGATATGAAGATCGAGACCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1921
Db 440 ArgGlnAspValArgAsnGlyAspProSerThrLeuCysSer----- 453
QY 1922 CACCATGGCCAGCGCCCTGAA-----GAGAGAATCATCTATGCTGTAGAGAT 1969
Db 454 -----GlyAspSerSerArgProAlaLeuLeuGluHisLysValPheGlyValGluGly 471
QY 1970 AGTAGCACATTTTGAATGCGAGTCGGAAGTCGAGAGAGCGCTGCTATTGGCAATTC 2029
Db 472 SerSerAlaPheLeuGluCysGluProArgSerLeuGlnAlaArgValGluTrpThrPhe 491
QY 2030 CAGAGCGGAATGAAGAGCGAAGAGAGATGATGATGATGATGATGATGATGATGATGAT 2089
Db 492 GlnArgAlaGlyValThrAlaHisThrGlnValLeuAlaGluGluArgThrGluArgThr 511
QY 2090 GATCAAGGCTCTGCTAGCTAGCTTACAAAGAGAGATTCAGGCAATTCCTCTGCCAT 2149
Db 512 AlaArgGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 531
QY 2150 GCGGTGAACATGGGTTCATCAAACTCTTCTTAAGGTAAACCCCTGGAAGTCAATGACACA 2209
Db 532 AlaValGluGlnGlyPheThrGlnProLeuArgLeuSerLeuHisValLeuSerAla 551
QY 2210 GAGCATTTGGAAGAATCTTCTCATAAAGATGATGATGATGATGATGATGATGATGAT 2269
Db 552 ThrGlnAlaGluArgLeu-----AlaArgAlaGluGlu 562
QY 2270 ATGTCCAATAGCATCACACTAGCCAGAGAGGTCTGTGATCAGAGACTTTCATGAGCTCATC 2329
Db 563 AlaAlaProAlaAlaProGlyProLysLeuTyrPheArgPheLeuGlnLeuVal 582
QY 2330 AACACCCCAATCTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2389
Db 582 ----- 582

QY 2390 AACACAACTCGGCAAGCCAGGACATACCCAGGAGACATGACAAATGGAAG----- 2443
Db 593 -----GluProGly-----GlyGlyGlySerAlaAsnSerLeuArgMetCys 596
QY 2444 -----CACTTACAAGAAAAATAAGAAAGGTAGAAACAGG 2476
Db 597 ArgProGlnProAlaLeuGlnSerLeuProLeuGluSerArgArgLysGlyArgAsnArg 616
QY 2477 AGGACCCAC-----GAATTTGAGAGGCCACCCAGGAGT 2509
Db 617 ArgThrHisAlaProGluProArgAlaGluArgGlyProArgSer 631
RESULT 9
QNS98 PRELIMINARY; PRT; 782 AA.
AC QNS98; Q9H7Q3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Senaphorin sem2 (FLJ00014 protein).
GN SEM2 OR FLJ00014.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Seki N., Hattori A., Hayaishi A., Kozuma S., Muramatsu M., Miyajima N.,
RA Saito T.;
RT "Human semaphorin";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 58-782 FROM N.A.
RC TISSUE=Spleen;
RA Ohara O., Nagase T., Kikuno R., Okumura K.;
RT "The nucleotide sequence of a long cDNA clone isolated from human
spleen";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB029496; BAA98132.1; -;
DR EMBL; AK024425; BAB15715.1; -;
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR001627; Sema.
DR Pfam; PF00047; IG; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS0835; IG_LIKE; 1.
SQ SEQUENCE 782 AA; 86700 MW; 85CB424874DF6663 CRC64;
Alignment Scores:
Pred. No.: 1.09e-146 Length: 782
Score: 1830.00 Matches: 366
Percent Similarity: 64.17% Conservative: 139
Best Local Similarity: 46.51% Mismatches: 244
Query Match: 37.38% Indels: 38
DB: 4 Gaps: 13
US-09-774-490-1 (1-2709) x QNS98 (1-782)
QY 221 GTCTGT---CTTTCTGGGAGTATTACTTACAGACAGCAAACTATCAGATGGGAG 277
Db 8 IleCysTrpLeuLeuGlyGlyLeuLeuHisGlyGlySerSerGlyProSerProGly 27
QY 278 AACATGTGCAAGCTGAATTTATCTTACAAAGAAATGTTGGAATCCAAATGTGATC 337
Db 28 ProSerValProArgLeuArgLeuSerTyrArgAspLeuLeuSerAlaAsnArgSerAla 47
QY 338 ACTTTCAATGGCTTGGCCAAACAGCTCCAGTTTCATATCATCTTCTTCTTGGATGAGAACGG 397
Db 338 ACTTTCAATGGCTTGGCCAAACAGCTCCAGTTTCATATCATCTTCTTCTTGGATGAGAACGG 397

Db 48 IlePheLeuGlyProGlnGlySerLeuAsnLeuGlnAlaMetTyrLeuAspGluTyrArg 67
QY 398 AGTAGCTGTATGTTGGAGCAAGAGTACATATTTTCATTCGACCTGGTTAATATC--- 454
Db 68 AspArgLeuPheLeuGlyGlyLeuAspAlaLeuTyrSerLeuArgLeuAspGlnAlaTrp 87
QY 455 AAGGATTTTCAAAAGATTGTGGCCAGTATCTTACACAGAGAGATGAATGCAAGTGG 514
Db 88 ProAspProArgGluValLeuTrpProGlnProGlyGlnArgGluGluCysValArg 107
QY 515 GCTGGAAGAACATCTGTAAGAAATGTCTTAATTTTCATCAAGTACTTAAAGGCATATAT 574
Db 108 LysGlyArgAspProLeuThrGluCysAlaAsnPheValArgValLeuGlnProHisAsn 127
QY 575 CAGACTCATTGTCAGCCCTGTGGAACGGGGCTTTTCATCCCAATTTGACCTACATTTGAA 634
Db 128 ArgThrHisLeuLeuAlaCysGlyThrGlyAlaPheGlnProThrCysAlaLeuIleThr 147
QY 635 ATTGGACATCATCTGAGGACAAATATTTTAAGCTGGGAGAACTCACATTTTGAACGGC 694
Db 148 ValGlyHisArgGlyGlu---HisValLeuHisLeuGluProGlySerValGluSerGly 166
QY 695 CTTGGAAGAGTCCATATACCTTAAGCTGCTGACAGCATCTCTTTTAATAGATGAGAA 754
Db 167 ArgGlyArgCysProHisGluProSerArgProPheAlaSerThrPheIleAspGlyGlu 186
QY 755 TTATCTCTGGAACCTGACCTGATTTTATGGGCGAGACTTTGCTATCTTCCGAACTCTT 814
Db 187 LeuTyrThrGlyLeuThrAlaAspPheLeuGlyArgGluAlaMetIlePheArgSerGly 206
QY 815 GGGCACCACCCCAATCAGGACAGACAGCATGATTCAGGTGCTCAATGATCAAGCA 874
Db 207 GlyProArgProAlaLeuArgSerAspSer---AspGlnSerLeuLeuHisAspProArg 225
QY 875 TTCTTAGTGGCCACCTCTCAGAGCTGACATCTCTGAAGATGACAAAGTATATCTTT 934
Db 226 PheValMetAlaAlaArgIleProGluAsnSerAspGlnAsnAspLysValTyrPhe 245
QY 935 TTCTCCGCGAAAT-----GCAATAGATGGAGAACACTCTCGAAAGACTACTCAGCT 988
Db 246 PhePheSerGluThrValProSerProAspGly---GlySerAsnHisValThrValSer 264
QY 989 AGAATAGTTCAGATGCAAGAAATGACTTTGGAGGGCAAGAGTCTGTGTGAATAAATGG 1048
Db 265 ArgValGlyArgValCysValAsnAspAlaGlyGlnArgValLeuValAsnLysTrp 284
QY 1049 ACAACATCTCCAAAGCTGCTGATTTGCTCAGTCCAGGCTCCAAATGGCATTCACACT 1108
Db 285 SerThrPheLeuLysAlaAlaArgLeuValCysSerValProGlyProGlyAlaGluThr 304
QY 1109 CATTTTGATGAATCAGGATGATTTCTTAATGAACTTTAAAGATCTCTAAATAATCCAGTT 1168
Db 305 HisPheAspGlnLeuGluAspValPheLeuLeuTrpProLysAlaGlyLysSerLeuGlu 324
QY 1169 GTATATGAGTGTTCAGCATCTCCAGTAACTTTCAAGGATCAGCGCTGTGTATGAT 1228
Db 325 ValTyrAlaLeuPheSerThrValSerAlaValPheGlnGlyPheAlaValCysValTyr 344
QY 1229 AGCATGAGTGTGAGAGGGTGTCTTGTGCTCATATGCCCAAGGATGGACCCAAAC 1288
Db 345 HisMetAlaAspIleTrpGluValPheAsnGlyProPheAlaHisArgAspGlyProGln 364
QY 1289 TATCAATGGGTCCCTTATCAAGGAAGAGTCCCTATCCAGGCCAGGAACTTGTCCACG 1348
Db 365 HisGlnTrpGlyProTyrGlyLysValProPheProArgProGlyValCysProSer 384
QY 1349 AAAACATTTGGT-----GGTTTGTACTTCAAAAGGACCTTCTGATGATGTT 1396
Db 385 LysMetThrAlaGlnProGlyArgProPheGlySerThrLysAspTyrProAspGluVal 404
QY 1397 ATAACCTTTGCAAGAGTATCCACCATGTACATCCAGTGTTCCTTATGAACATCGC 1456
Db 405 LeuGlnPheAlaAlaHisProLeuMetPheTrpProValArgProArgHisGlyArg 424

QY 1457 CCAATAGTATCAAAACGAGTCTAAATTTATCAATTTACAAATTTGCTAGACCCAGTG 1516
Db 425 ProValLeuValLysThrHisLeuAlaGlnGlnLeuHisGlnIleValValAspArgVal 444
QY 1517 GATCAGAGATGGACAGTATGATGTTATGTTTATCGGAACAGATGTTGGGACCGTTCTT 1576
Db 445 GluAlaGluAspGlyThrTyrAspValIlePheLeuGlyThrAspSerGlySerValLeu 464
QY 1577 AAAGTAGTTTCAATTTCTTAAAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGAAGA 1636
Db 465 LysValIleAlaLeuGlnAlaGlyGlySerAlaGluProGluGluValValLeuGluGlu 484
QY 1637 ATGACAGTTTTCGGGAACCGACTCTATTTCAGCAATGGAGCTTTCCACTAAGCAGCAA 1696
Db 485 LeuGlnValPheLysValProThrProIleThrGluMetGluIleSerValLysArgGln 504
QY 1697 CAATATATATTTGTTCAACGGCTGGGTTCGCCAGCTCCCTTTTACCGGTGTGATATT 1756
Db 505 MetLeuTyrValGlySerArgLeuGlyValAlaGlnLeuArgLeuHisGlnCysGluThr 524
QY 1757 TACGGAAAGCGTGTGCTGAGTGTTCCTCCCGAGAGCCCTTACTGTGCTGGATGGT 1816
Db 525 TyrGlyThrAlaCysAlaGluCysLeuAlaArgAspProTyrCysAlaTrpAspGly 544
QY 1817 TCTGATCTTCTCGCTATTTTCCCACT---GCAAGAGAGCGCACAAAGACCAAGATATA 1873
Db 545 AlaSerCysThrHisTyrArgProSerLeuGlyLysArgA-GpPheArgArgGlnAspIle 564
QY 1874 AGAATGAGAGCCCACTGACTCCTGTTTACAGCTTTACACCATGATTAATCAACATGCCAC 1933
Db 565 ArgHisGlyAsnProAlaLeuGlnCysLeu-----GlyGln 576
QY 1934 AGCCTGGAAGAGAGA-----ATCATCTATGCTGTAGAGAAAT 1969
Db 577 SerGlnGluGluGluAlaValGlyLeuValAlaAlaThrMetValTyrGlyThrGluHis 596
QY 1970 AGTAGCATTATTTGGAAATGCGAGTCCGAAGTCCGAGAGCGCTGCTTATTGGCAATTC 2029
Db 597 AsnSerThrPheLeuGluCysLeuProLysSerProGlnAlaAlaValArgTrpLeuLeu 616
QY 2030 CAGAGCGCAATTAAGAGCGGAAAGAGATCAGAGTGGATGATCATATCATCAGGACA 2089
Db 617 GlnArgProGlyAspGluGlyProAspGlnValLysThrAspGluArgValLeuHisThr 636
QY 2090 GATCAAGCCTTCTGCTAGTGTACTACAACAGAGAGGATTCAGGCAATTACTCTGCCAT 2149
Db 637 GluArgGlyLeuLeuPheArgLeuSerArgPheAspAlaGlyThrTyrThrCysThr 656
QY 2150 CGCGTGAACATGGTTTCATAAACTCTTAAAGGTAAACCTGGAAGTCAATTCACACA 2209
Db 657 ThrLeuGluHisGlyPheSerGlnThrValValArgLeuAlaLeuValValIleValAla 676
QY 2210 GAGCATTGGAAGAACTTCTTATAAAGATGATGAGAGTGGCTCTTAAACCCAAAGAA 2269
Db 677 SerGlnLeuAspAsnLeuPheProGluProLysProGluProAlaArgGly 696
QY 2270 ATGTCCAATAGCATGACACCTTAGCCAGAGTCTGTTACAGAGACTTCATCAGCTCATC 2329
Db 697 GlyLeuAlaSerThrProPro-----LysAlaTrpTyrLysAspIleLeuGlnLeuIle 714
QY 2330 AACCACCCCAATCTCAACAGATGGATGAGTTCGTGAACAAGTTTGGMAAGGACCGCA 2389
Db 715 GlyPheAlaAsnLeuProArgValAspGluTyrCysGluArgValTrpCysArgGlyThr 734
QY 2390 AAACAA-----CGTCGCAAGCGCCAGACATACCCAGGGAACAGT 2431
Db 735 ThrGluCysSerGlyCysPheArgSerArgGlyLysGlnAlaArgGlyLysSer 754
QY 2432 AACAAATGGAAGCACTTACAGAAAAATAGAAAGGTAGAAACAGGAGACCCAGCAATTT 2491
Db 755 -----TrpAlaGlyLeuGluLeuGlyLysLysMetLysSerArgValHisAlaGluHis 772

QY 2492 GAGAGGGCACCCAGGAGTGTCT 2512
 DB 773 AsnArgThrProArgGluVal 779
 RESULT 10
 Q8NCQ1 PRELIMINARY; PRT; 416 AA.
 AC Q8NCQ1;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue=Testis;
 RA Strausberg R.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC029590; AAH29590.1; -;
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR003659; Plectin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF00447; IG; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 416 AA; 48691 MW; 61D92D89CAE70631 CRC64;
 Alignment Scores:
 Pred. No.: 1,84e-96 Length: 416
 Score: 1238.50 Matches: 227
 Percent Similarity: 72.73% Conservatives: 85
 Best Local Similarity: 52.91% Mismatches: 100
 Query Match: 25.30% Indels: 17
 DB: 4 Gaps: 8
 US-09-774-490-1 (1-2709) x Q8NCQ1 (1-416)
 QY 1232 ATGAGTGATGTGAGAGGGTGTCTTGTGTCATATGCCAGGATGCCAGGATGCCAAGTAT 1291
 DB 1 MetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysGluSerAlaAspHis 20
 QY 1292 CAATGGTCCCTTATCAAGGAAGAGTCCCTTATCCAGGCCAGGAACTTGTCCAGCAAA 1351
 DB 21 ArgTrpValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCysProSerLys 40
 QY 1352 ACATTGGTGGT---TTTGACTCTACAAAGGACCTTCTCTGATGATGTTATTAACCTTTGCA 1408
 DB 41 ThrTyrAspProLeuIleLysSerThrArgAspPheProAspAspValIleSerPheIle 60
 QY 1409 AGAAGTCCAGCCATGACATCCAGTGTCTTATGACAAATCGCCCAATGATGATC 1468
 DB 61 LysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyProThrPheLys 80
 QY 1469 AAAACGGGTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1528
 DB 81 ArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIleAlaGluAsp 100
 QY 1529 GGACAGTATGATGTTATGTTATCGGACAGATGTTGGACCGTCTCTTAAGTAGTTTCA 1588
 DB 101 GlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLysValValSer 120
 QY 1589 ATTCCTAAGGAGACTGGTATGATTTAGAGAGGTTCTCTGGAAGAATGACAGTTT 1648
 DB 121 IleSerLysGlyLysTrp---AsnMetGluGluValValLeuGluGluLeuGlnIlePhe 139

QY 1649 CGGAAACCGACTGCTATTTTTCAGCAATGAGGCTTCTTCCACTAAGACCAACAATATATTT 1708
 DB 140 LysHisSerSerIleIleLeuAsnMetGluLeuSerLeuLysGlnGlnLeuTyrIle 159
 QY 1709 GGTTCACAGGCTGGGTGCTCCAGTCTCCTTTACACCGGTGTGATATTTACGGGAAGCG 1768
 DB 160 GlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyrGlyLysAla 179
 QY 1769 TGTCTGAGTGTCTGCTCGCCGAGACCTTACTGTCTGCTGGATGTTCTGTCATGTTCT 1828
 DB 180 CysAlaAspCysCysLeuAlaAspProTyrCysAlaIlePheValTyrGlyAspPro 199
 QY 1829 CGCTATTTCCTCACTGCAAGAGACGACCAAGACGACCAAGATATAAGAAATGAGACCA 1888
 DB 200 ArgTyrAlaProThrSerLysArgAlaArgGlnAspValLysTyrGlyAspPro 219
 QY 1889 CTGACTCCTGCTTACACCATGATATCACCATGCCACACCCCTGGAAGAGAGA 1948
 DB 220 IleThrGlnCysTrpAspIleGluAspSerIleSerHis---GluThrAlaAspGluLys 238
 QY 1949 ATCATCTATGCTAGAGATAGTAGACATTTTGGATGCTCAGTCCGAGTCCGAGAGA 2008
 DB 239 ValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLysSerGlnGln 258
 QY 2009 CGCTGGTCTATTGGCAATTCACAGAGCGCAATGAAGAGCGCAAGAGAGATCAGAGTG 2068
 DB 259 AlaThrIleLysTrpTyrIleGlnArgSerGlyAspGluHisArgGluGluLysPro 278
 QY 2069 GATGATCATATCATCAGACAGATCAAGCCCTTCTGCTACGTAGTCTTACACAGAGAT 2128
 DB 279 AspGluArgIleIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGlnLysAsp 298
 QY 2129 TCAGGCAATTACCTCTGCCATCGGTGGAACATCGGTTCATCAAACTCTTCTTAAGTGA 2188
 DB 299 SerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIleValLysLeu 318
 QY 2189 ACCCTGGAAGTCATTGACACAGAGCATTTGGAAGAACTTCTTCAAAAGATGATGGA 2248
 DB 319 ThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGluHis---- 336
 QY 2249 GATGGCTTAAAGACCAAGAAATGTCATAGCATGACACCTAGCAGCAAGGTCTGGTAC 2308
 DB 337 GluGluGlyLysValLysAsp-----LeuLeuAlaGluSerArgLeuArgTyr 352
 QY 2309 AGAGACTTCATGACGCTCATCAACACCCCAATCTCAACACGATGATGATGATGTCGAA 2368
 DB 353 LysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGlnTyrCysGlu 371
 QY 2369 CAAGTTTGGAAAGGGAGCCGAAACACACGTCGGCAAGGCCAGGACATACCCAGGGAAC 2428
 DB 372 GlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly----- 386
 QY 2429 AGTAACAAATGGAAGCATTACAAGAAATTAAGAAAGTAGAAGACAGAGAGCCAC--- 2485
 DB 387 ---ProLysTrpLysHisMetGlnGluMetLysLysLysArgAsnArgArgHisArg 405
 QY 2486 GAAATTGAGGGGACCCAGGAGTGC 2512
 DB 406 AspLeuAspGluLeuProArgAlaVal 414
 RESULT 11
 Q9HBR1 PRELIMINARY; PRT; 457 AA.
 ID Q9HBR1
 AC Q9HBR1;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]

Db	246	PheArgGluThrAlaValGluAlaProAlaLeuGlyArgLeuSerValSerArgVal	265
Qy	995	GGTCAGATATGCAAGAAATGACTTTTGGAGGCGCAGAGAAGTCTGGTGAATAAATGGCAACA	1054
Db	266	GlyGlnIleCysArgAsnAspValGlyGlyGlnArgSerLeuValAsnIleTyrThr	285
Qy	1055	TTCCCTCAAAGCTCGTCTGATTTGCTCAGTGCCAGGTCCAAATGGCAATGGACACTCATTTT	1114
Db	286	PheLeuLysAlaArgLeuValCysSerValProGlyValGluGly--AspThrHisPhe	304
Qy	1115	GATGAAGTCCAGGATGATATTCCTAATGAACCTTTAAAGATCCTTAAATAATCCAGTTGTATAT	1174
Db	305	AspGlnLeuGlnAspValPheLeuLeuSerSerArgAspHisArgThrProLeuLeuTyr	324
Qy	1175	GGAGTGTTCACGACTTCCAGTAACATTTTCAAGGGATCAGCGGTGTGTATGTATAGCATG	1234
Db	325	AlaValPheSerThrSerSerIlePheGlnGlySerAlaValCysValTyrSerMet	344
Qy	1235	AGTGATGTGAGAAGGTGTTTCTTGTGTCCATATGCCACAGGATGGACCCAACTATCAA	1294
Db	345	AsnAspValArgAlaPheLeuGlyProPheAlaHisLysGluGlyProMetHisGln	364
Qy	1295	TGGTGCTCTATCAAGAGAGAGTCCCTATTCACGCCGAGGAACCTGTCCAGCAAAACA	1354
Db	365	TyrValSerTyrGlnGlyArgValProTyrProArg-GlnAlaCysAlaProAlaArgPr	384
Qy	1355	TTTGGTGGTTTGACTCTACAAGGAGCCTTCTCGATGATGTATTAACCTTTGCAAGAAGT	1414
Db	384	oLeuAlaProSerValProProArgThrSerGlnThrMetSerSerLeuArgGlyTh	404
Qy	1415	CATCCAGCCATGT---ACAATCCAGTGT 1439	
Db	404	rThrProSerCysThrLeuSerCys 413	
RESULT 12			
Q8BKQ6			
ID	Q8BKQ6	PRELIMINARY;	PRT; 403 AA.
AC	Q8BKQ6;		
DT	01-MAR-2003 (TrEMBLrel. 23, Created)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)		
DE	Semaphorin 3D precursor homolog.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RP	[1]		
RN	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Dorsal root ganglion;		
RX	MEDLINE=22354683; PubMed=12466851;		
RA	The FANTOM Consortium,		
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;		
RT	"Analysis of the mouse transcriptome based on functional annotation of		
RT	60,770 full-length cDNAs."		
RL	Nature 420:563-573(2002).		
DR	ENBL: AK051165; BAC34542.1; -.		
SQ	SEQUENCE 403 AA; 45575 MW; 1D5B284549825884 CRC64;		
Alignment Scores:			
Pred. No.:	4.8e-88	Length:	403
Score:	1139.50	Matches:	216
Percent Similarity:	71.71%	Conservative:	73
Best Local Similarity:	53.60%	Mismatches:	100
Query Match:	23.27%	Indels:	14
DB:	11	Gaps:	5
US-09-774-490-1 (1-2709) x Q8BKQ6 (1-403)			
Qy	148	ACAATACAGGAGGAAGACTAAGACGACGAAGAGGACCTACAGGTCTGCAGCATGGGCTG	207
Db	4	ThrLysAspGluAsnProArgSerGlnAspLeu---HisLeuPheHis-AlaTr	22

```
QY 208 GTTAACAGGATGTC---TGCTTTTCTGGGAGATTACTTACAGCAAGCAAACTA 264
Db 22 pMetMetLeuIleMetThrValLeuPheLeuProValThrGluThrSer----- 38
QY 265 TCAGAAATGGGAAGAACAAATGTGCCAAGGTGAATATCTCTACAAAGAAATGTTGAATC 324
Db 39 -----LysGlnAsnIleProArgLeuLysLeuThrTyLysAspLeuLeuSe 55
QY 325 CAACATGTGATCATTCTCAATGGCTTGGCCACAGCTCCAGTTCATCATCTCTTTT 384
Db 55 rAsnThrCysIleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeu 75
QY 385 GGATGAGCAACGAGTAGCTGTATCTGGAGCAAGGATCATATTTTCATTCGACCT 444
Db 75 uaspGluGluArgGlyLeuLeuLeuGlyAlaLysAspHisValPheLeuLeuSerLe 95
QY 445 GGTTAATATC---AGGATTTTCAAAGATTGTGGCCAGTATCTTACACCAAGAGAA 501
Db 95 uValAspLeuAsnLysAsnPheLysLysIleTyTrpProAlaAlaLysGluArgValG 115
QY 502 TGAATGCAAGTGGCTGGAAGACATCTGGAAGATGTGCTAATTTTCATCAAGTACT 561
Db 115 uLeuCysLysLeuAlaGlyLysAspAlaAsnAlaGluCysAlaAsnPheIleArgVal 135
QY 562 TAAGGCATATATACAGACTCACTGTATCGCTGTGGAACGGGGCTTTTCATCCAATTG 621
Db 135 uGlnProTyAsnLysThrHisValTyValCysGlyThrGlyAlaPheHisProLeuCy 155
QY 622 CACCTACATTAATTTGACATCATCTCGAGCAATATTTTAAAGCTGGAGAACTCACA 681
Db 155 sGlyTyrlleAspLeuGlyAlaAsnLysGluLeuIlePheLysLeuAspThrHisAs 175
QY 682 TTTTGAAGCGCGTGGGAAGATCCATATACCTTAAGCTGCTGACAGCATCTTTT 741
Db 175 nLeuGluSerGlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerVal 195
QY 742 AATAGATGAGAATATATCTCGAACTCAGCTGATTTTATGGCGGAGACTTTGCTAT 801
Db 195 tThrAspGluHisLeuTySerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPh 215
QY 802 CTTCCGAATCTTGGG-----CACCAACCAATCATCAGGACAGACAGCATGATTC 852
Db 215 eThrArgSerLeuGlyLeuMetGlnAspHisSerIleArgThrAspIleSerGluHi 235
QY 853 CAGGTGGCTCAATGATCCAAAGTTCATTAGTCCCACTCTCAGAGAGTGAACATCC 912
Db 235 shistrPLeuAsnGlyAlaLysPheIleGlyThrPheProIleProAspThrTyAsnPr 255
QY 913 TGAAGATGACAAAGTATATCTTTCTTCGCGTGAATGCAATAGATGAGCAACTCTGG 972
Db 255 oAspAspAspLysIleTyPhePhePheArgGluSerSerGlnGlySerThrSerAs 275
QY 973 AAAAGCTACTCAGTAGAATAGTTCAGATATGCAAGAATGACTTTGGAGGCGACAGAG 1032
Db 275 pArgSerIleLeuSerArgValGlyArgValCysLysAsnAspValGlyGlnArgSe 295
QY 1033 TCTGGTGAATAATGACCAACTCTCAAGCTGCTGATGTTCTGCTAGTCCAGGTCC 1092
Db 295 rLeuIleAsnLysTrpThrPheLeuLysAlaArgLeuIleCysSerIleProGlySe 315
QY 1093 AAATGGCATGTGACTCATTTTGTATGAACTGAGATGATTTCTTAATGAATTTAAAGA 1152
Db 315 rAspGlyAlaAspThrHisPheAspGluLeuGlnAspIleTyLeuLeuProThrArgAs 335
QY 1153 TCCTAAAATCCAGTTGTATATGGAGTGTATTACGACTTCCAGTACATTTTCAAGGATC 1212
Db 335 pGluArgAsnProValValTyGlyValPheThrIysThrSerSerIlePheLysGlySe 355
QY 1213 ACCCGTGTATGTATAGCATGATGTGAGAGGGGTTCCTTGTGTCATATGCCCA 1272
Db 355 rAlaValCysValTySerMetAlaAspIleAlaArgAlaValPheAsnGlyProTyAlaHi 375
QY 1273 CAGGGATGACCCCACTATCAATGGTGTGCTTATCAAGGAAGAGTCCCTATCCAGGCC 1332
```

```
Db 375 sLysGluSerAlaAspHisArgTrpValGlnTyAspGlyArgIleProTyProArgPr 395
QY 1333 AGGAAC 1339
Db 395 oGlyThr 397
RESULT 13
Q9J129 PRELIMINARY; PRT; 296 AA.
ID Q9J129
AC Q9J129;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Semaphorin (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RA Brenz Verca M.S., Widmer D.A.J., Wagner G.C., Dreyer J.L.;
RT "New rat semaphorin with closest homology to avian Sema3D (C-
collapsin-2).";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF268594; AAF76329.1; -.
DR InterPro; IPR001627; Sema.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00630; Sema; 1.
FT NON_TER 1
FT NON_TER 296
SQ SEQUENCE 296 AA; 33558 MW; C2226212885D044C CRC64;
Alignment Scores:
Pred. No.: 5,08e-83 Length: 296
Score: 1080.00 Matches: 192
Percent Similarity: 81.82% Conservative: 51
Best Local Similarity: 64.65% Mismatches: 52
Query Match: 22.06% Indels: 2
DB: Gaps: 2
US-09-774-490-1 (1-2709) x Q9J129 (1-296)
QY 914 GAAGATGCAAGATATCTTTTCTCGTGAATAATGCAATAGTAGGAGAACACTCTGGA 973
Db 1 AspAspAspLysIleTyPhePhePheArgGluSerGlnGlySerThrSerAsp 20
QY 974 AAAGCTACTCAGCTAGTAATAGTGCAGATATGCAAGATGACTTTGGAGGCGACAGAGT 1033
Db 21 ArgSerIleLeuSerArgValGlyArgValCysLysAsnAspValGlyGlnArgSer 40
QY 1034 CTGGTGAATAATGGACAACTTCTCCTCAAGCTGCTGATGTTGCTAGTCCAGGTCCA 1093
Db 41 LeuIleAsnLysTrpThrPheLeuLysAlaArgLeuValCysSerIleProGlySer 60
QY 1094 AATGCGATTGACATCTTTTGTATGCACTGAGGATGTTCTTAATGAACCTTTAAAGAT 1153
Db 61 AspGlyAlaAspThrHisPheAspGluLeuGlnAspIleTyLeuLeuProThrArgAsp 80
QY 1154 CTTAAAATCCAGTTGTATATGAGTGTATACGACTTCCAGTAACTTTTCAAGGGATCA 1213
Db 81 GluArgAsnProValValTyGlyValPheThrThrSerSerIlePheLysGlySer 100
QY 1214 CCGTGTGTATGTATAGCATGATGTGAGAGGGGTTCCTTGTGTCATATGCCCA 1273
Db 101 AlaValCysValTySerIleAlaAspIleAlaValPheAsnGlyProTyAlaHis 120
QY 1274 AGGGATGGACCCCACTATCAATGGTGTGCTTATCAAGGAAGAGTCCCTATCCAGGCCA 1333
Db 121 LysGluSerAlaAspHisArgTrpValGlnTyAspGlyArgIleProTyProArgPr 140
QY 1334 GGAACTTGTCCCAAGCAAAACATTTGGTGGT---TTTGTACTACAAAGGACCTTCTCTGAT 1390
```

```
Db 141 GlyThrCysProSerLysThrTyrAspProLeuIleLysSerThrArgAspPheProAsp 160
|||||
QY 1391 GATGTTAAACCTTTTGCACAAAGTCCATCCAGCATGTAACATCCAGTGTTCCTATGAAC 1450
|||||
Db 161 AspValIleSerPheIleArgHisProValMetPheLysSerValTyrProValAla 180
|||||
QY 1451 AATCCCAATAGTGAATCAACAGGATGTAATATCAATTTACACAAATTTGCTGAGAC 1510
|||||
Db 181 GlyAlaProThrPheGlnArgIleAsnValAspTyrArgLeuThrGlnIleValValAsp 200
|||||
QY 1511 CGAGTGGATGAGAGATGCGAGATGATGTTATGTTATCCGACAGATGTTGGGACC 1570
|||||
Db 201 HisValValAlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAsp***GlyThr 220
|||||
QY 1571 GTTCTTAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAGAGGTTCTGCTG 1630
|||||
Db 221 ValLeuLysValValSerIleSerLysGluLysTrp---AsnMetGluGluValValLeu 239
|||||
QY 1631 GAAGAAATGACAGTTTTCGGGAACCGACTGCTATTTACGAAATGGAGCTTTCACACTAAG 1690
|||||
Db 240 GluGluLeuGlnValPheLysHisProThrAlaIleLeuAsnMetGluLeuSerLeuLys 259
|||||
QY 1691 CAGCAACATATATATTGGTTCAACGCTGGGTTGCCAGCTCCCTTTACACCGGTGT 1750
|||||
Db 260 GlnGlnGlnLeuTyrValGlySerTrpAspGlyLeuValGlnLeuSerLeuHisArgCys 279
|||||
QY 1751 GATATTTACGGGAACCGTGTGCTGAGTGTGCTGCTGCCCGCAGAGACCTTTAC 1801
|||||
Db 280 AspThrTyrGlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyr 296
|||||
```

RESULT 14

Q8BJC1 ID Q8BJC1 PRELIMINARY; PRT; 799 AA.

AC Q8BJC1; DT 01-WAR-2003 (TRENBLrel. 23, Created)

DT 01-WAR-2003 (TRENBLrel. 23, Last sequence update)

DT 01-WAR-2003 (TRENBLrel. 23, Last annotation update)

DE Senaphorin 4D precursor (Fragment).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_taxid=10090;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=NOD; TISSUE=Thymus;

RX MEDLINE=22354683; PubMed=12466851;

RA The RIKEN Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs."

RL Nature 420:563-573(2002).

DR EMBL; AK088653; BAC40480.1; --

FT NON TER 1

SQ SEQUENCE 799 AA; 88814 MW; 415D3B687150A59A CRC64;

Alignment Scores:

Pred. No.:	4.18e-74	Length:	799
Score:	976.50	Matches:	253
Percent Similarity:	53.53%	Conservative:	111
Best Local Similarity:	37.21%	Mismatches:	238
Query Match:	19.94%	Indels:	78
DB:	11	Gaps:	27

US-09-774-490-1 (1-2709) x Q8BJC1 (1-799)

QY 407 TATGTTGGCAAGATGACATATTTTCATCCAGCTGGTAAATATCAAGATTTTCAA 466

Db 1 TyrValGlyAlaArgGluAlaValPheAlaValAsnAlaLeuAsnIleSerGluLysGln 20

QY 467 ---AAGATTGTGCGCCAGTATCTTACACAGAGAGATGAATCGAAGTGGCTGGAAA 523

Db 21 HisGluValTyrTrpLysValSerGluAspLysSerLysCysAlaGluLysGlyLys 40

```
QY 524 GACATCTCGAAGAATGTCTTAATTTTCATCAAGGTACTTAAGGCATATAATCAGACTCAC 583
|||||
Db 41 SerLysGlnThrGluCysLeuAsnTyrIleArgValLeuGlnProLeuSerArgThrSer 60
|||||
QY 584 TTGTACCGCTGTGAACGGGGCTTTTCATCCAAATTTCCACCTACATTAATGAATGGACAT 643
|||||
Db 61 LeuTyrValCysGlyThrAsnAlaPheGlnProThrCysAspHisLeuAsnLeu----- 78
|||||
QY 644 CATCTGAGGACAATATTTTAAAGCTGGAGAACTCACATTTTGAACCGCGCTGGGAAG 703
|||||
Db 79 -----ThrSerPheLysPheLeuGlyLysSer---GluAspGlyLysGlyArg 93
|||||
QY 704 AGTCCATATGACCTTAAGCTGCTCAGACATCCCTTTTAAATAGATGGAGAATTATACCT 763
|||||
Db 94 CysProPheAspProAlaHisSerTyrThrSerValMetValGlyGlyGluLeuTyrSer 113
|||||
QY 764 GGAACCTGAGCTGATTTATGGGGCGAGACTTCTCTATCTCCGAACTCTTGGGGCACCAC 823
|||||
Db 114 GlyThrSerTyrAsnPheLeuGlySerGluProIleLeuSerArgAsn---SerSerHis 132
|||||
QY 824 CACCAATCAGACAGACAGCAGCATGATTCACAGGTGGCTCAATGATCCAAAGTTCAATTAGT 883
|||||
Db 133 SerProLeuArgThrGlu---TyrAlaIleProIleLeuAsnGluProSerPheValPhe 151
|||||
QY 884 GCCCACCCTCATCTCAGAGAGT---GACAATCCT-----GAAGATGACAAAGTATACCTT 934
|||||
Db 152 AlaAspValIleGlnLysSerProAspGlyProGluGlyGluAspAspLysValTyrPhe 171
|||||
QY 935 TTCTTCGCTGAAATGCAATAGATAGGAGACACTCTGGAAAAGCTACTCAGCTAGATA 994
|||||
Db 172 PhePheThrGluValSerValGluTyrGluPheValPheLysLeuMetIleProArgVal 191
|||||
QY 995 GGTGAGATATGCAAGAATGACTTTGGAGGGGCACAGAGTCTGCTGAATAATGACACACA 1054
|||||
Db 192 AlaArgValCysLysGlyAspGlnGlyLeuArgThrLeuGlnLysLysTrpThrSer 211
|||||
QY 1055 TTCTTCAAGCTCGTCTGATTTGCTCAGTGTCCAGTGTCCAAATGCGCATGACACATTTT 1114
|||||
Db 212 PheLeuLysAlaArgLeuIleCysSerLysProAspSerGlyLeuVal-----Phe 228
|||||
QY 1115 GATGAATGCGAGATGATTTCTTAATGAATCTTTAAAGATCTCTAAATCCAGTTGTATAT 1174
|||||
Db 229 AsnIleLeuGlnAspValPheValLeuArgAlaProGlyLeuLysGluProValPheTyr 248
|||||
QY 1175 GGAGTGTTCAGACTTCCAGTAACATTTTCAAGGATCAGCGGTGTGTATGTATAGCATG 1234
|||||
Db 249 AlaValPheThrProGlnLeuAsnAsnValGlyLeuSerAlaValCysAlaTyrThrLeu 268
|||||
QY 1235 AGTGAATGAGAGGGTGTTC---CTTGGTCCATAT-----GCCACAGGATGGA 1282
|||||
Db 269 AlaThrValGluAlaValPheSerArgGlyLysTyrMetGlnSerAlaThrValGluGln 288
|||||
QY 1283 CCCAATCATCAATGGGTGCTTATCAAGGAAGATGCCCTATCCAGCGCCAGGACTTGT 1342
|||||
Db 289 SerHisThrLysTrpValArgTyrAsnGlyProValProThrProArgProGlyAlaCys 308
|||||
QY 1343 -----CCCACAAAACATTTGGTGGTGTTCGACTCTACAAAGGACCTTCTGTATGATGTT 1396
|||||
Db 309 IleAspSerGluAlaArgAlaAlaAsnTyrThrSerSerLeuAsnLeuProAspLysThr 328
|||||
QY 1397 ATAACCTTTGCAAGAAGTCCATCCAGCATGTACATCCAGTGTTCCTATCAACAAATCGC 1456
|||||
Db 329 LeuGlnPheValLysAspHisProLeuMetAspAspSerValThrProIleAspAsnArg 348
|||||
QY 1457 CCAATAGTATCAAAACGGATGTAATATCAATTTACAAATTTGCTGAGACCGATG 1516
|||||
Db 349 ProlLysLeuIleLysLysAspValAsnTyr-----ThrGlnIleValValAspArgThr 366
|||||
QY 1517 GATCAGAAATGACAGACAG---TATGATGTTATGTTTATCGAACAGATGTTGGACCGTT 1573
|||||
Db 367 GlnAlaLeuAspGlyThrPheTyrAspValMetPheIleSerThrAspArgGlyAlaLeu 386
|||||
```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 10:02:39 ; Search time 63 Seconds
(without alignments)
10213.340 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 4896

Sequence: 1 aatcttttattttatcgatg.....aggctttttttcttaataacc 2709

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 451899 seqs, 118759770 residues

Total number of hits satisfying chosen parameters: 903798

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-O=/cgn2_1/USPTO_spool/US09774490/runat_03082003_095418_25421/app_query.fasta_1.2887
-DB=Published Applications AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNIT8=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MOD=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09774490 @CGN 1 1 85 @runat_03082003_095418_25421
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
17:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	4201	85.8	771	15	US-10-097-340-284	Sequence 284, App
2	4201	85.8	771	15	US-10-262-538-10	Sequence 10, Appl
3	2245.5	45.9	777	11	US-09-946-374-310	Sequence 310, App
4	2245.5	45.9	777	14	US-10-052-586-348	Sequence 348, App
5	2245.5	45.9	777	15	US-10-174-590-348	Sequence 348, App
6	2245.5	45.9	777	15	US-10-176-758-348	Sequence 348, App
7	2245.5	45.9	777	15	US-10-175-737-348	Sequence 348, App
8	2245.5	45.9	777	15	US-10-173-706-348	Sequence 348, App
9	2245.5	45.9	777	15	US-10-175-738-348	Sequence 348, App
10	2245.5	45.9	777	15	US-10-175-752-348	Sequence 348, App
11	2245.5	45.9	777	15	US-10-176-482-348	Sequence 348, App
12	2245.5	45.9	777	15	US-10-176-757-348	Sequence 348, App
13	2245.5	45.9	777	15	US-10-176-913-348	Sequence 348, App
14	2245.5	45.9	777	15	US-10-180-552-348	Sequence 348, App
15	2245.5	45.9	777	15	US-10-180-557-348	Sequence 348, App
16	2245.5	45.9	777	15	US-10-173-700-348	Sequence 348, App
17	2245.5	45.9	777	15	US-10-174-572-348	Sequence 348, App
18	2245.5	45.9	777	15	US-10-174-579-348	Sequence 348, App
19	2245.5	45.9	777	15	US-10-174-588-348	Sequence 348, App
20	2245.5	45.9	777	15	US-10-175-739-348	Sequence 348, App
21	2245.5	45.9	777	15	US-10-175-740-348	Sequence 348, App
22	2245.5	45.9	777	15	US-10-175-743-348	Sequence 348, App
23	2245.5	45.9	777	15	US-10-176-488-348	Sequence 348, App
24	2245.5	45.9	777	15	US-10-176-492-348	Sequence 348, App
25	2245.5	45.9	777	15	US-10-176-747-348	Sequence 348, App
26	2245.5	45.9	777	15	US-10-176-750-348	Sequence 348, App
27	2245.5	45.9	777	15	US-10-176-985-348	Sequence 348, App
28	2245.5	45.9	777	15	US-10-176-987-348	Sequence 348, App
29	2245.5	45.9	777	15	US-10-176-992-348	Sequence 348, App
30	2245.5	45.9	777	15	US-10-176-993-348	Sequence 348, App
31	2245.5	45.9	777	15	US-10-184-658-348	Sequence 348, App
32	2245.5	45.9	777	15	US-10-173-695-348	Sequence 348, App
33	2245.5	45.9	777	15	US-10-173-697-348	Sequence 348, App
34	2245.5	45.9	777	15	US-10-173-705-348	Sequence 348, App
35	2245.5	45.9	777	15	US-10-174-576-348	Sequence 348, App
36	2245.5	45.9	777	15	US-10-174-585-348	Sequence 348, App
37	2245.5	45.9	777	15	US-10-174-586-348	Sequence 348, App
38	2245.5	45.9	777	15	US-10-175-747-348	Sequence 348, App
39	2245.5	45.9	777	15	US-10-175-748-348	Sequence 348, App
40	2245.5	45.9	777	15	US-10-176-481-348	Sequence 348, App
41	2245.5	45.9	777	15	US-10-176-485-348	Sequence 348, App
42	2245.5	45.9	777	15	US-10-176-487-348	Sequence 348, App
43	2245.5	45.9	777	15	US-10-176-493-348	Sequence 348, App
44	2245.5	45.9	777	15	US-10-176-493-348	Sequence 348, App
45	2245.5	45.9	777	15	US-10-176-756-348	Sequence 348, App

ALIGNMENTS

RESULT 1

US-10-097-340-284
; Sequence 284, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNANAVARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KANATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030

```

; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 284
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-284

```

Alignment Scores:

```

Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 15 Gaps: 0

```

US-09-774-490-1 (1-2709) x US-10-097-340-284 (1-771)

```

QY 200 ATGGGCTGTTAACTAGGATTTCTCTTTCTGGGAGATTAATCTACAGCAAGACA 259
DB 1 MetGlyThrLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGAAATGGGAAGCAATGTGCCAAGCTGAAATTAATCTACAAAGAAATGTTG 319
DB 21 AsnTyrGlnAsnGlyLysAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAACTATGTATCACTTTCAATGCTTGGCCCAAGCTCCAGTTATCATACCTTC 379
DB 41 GluSerAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
QY 380 CTTTGGATCAGGAACGAGTAGCTGTATGTTGGAGCAAGGATCACATATTTTCATTC 439
DB 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAATCAAGGATTTTCAAGATTTGTCGCCAGTATCTTACACCAAGACA 499
DB 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATGAATCAAGTGGCTCGAAAGACATCCCTGAAAGATGCTTAATTTTCATCAAGGTA 559
DB 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCATTGTACGCTGTGGAACGGGGCTTTTCATCCAAAT 619
DB 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTGAATTTGGACATCATCTCTGAGGACAATATTTTAAGCTGGAGAACTCA 679
DB 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAAGCGCGTGGAGAGTCCATATGACCTTAAGCTGTGACAGCATCCCTT 739
DB 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGGAATTAATCTCTGGAAGTGCAGCTGATTTTATGGGCGAGACTTTGCT 799
DB 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200

```

```

QY 800 ATCTTCGAACCTCTTGGSCACACCCACCAATCAGGACAGAGCAGCATGATTTCCAGGTGG 859
DB 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCAATGATCCAAAGTTTCATTAGTGGCCACCTCATCTCAGAGAGTGACAATCTCTGAAGAT 919
DB 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTTCTTCCTGGTGAATAATGCAATAGATGGAGAACACTCTGGAAAAGCT 979
DB 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGATAGGTGCAGATATGCAAGATGACCTTTGGAGGSCACAGAAAGTCTGGTG 1039
DB 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGACAACATCTCTCAAGCTCGTCTGATTTGCTCAGTGCCAGGTCCAAATGCG 1099
DB 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTATTTTGATGAATCTGAGGATGTATTCCTAATGAATTTAAAGATCCTAAA 1159
DB 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGTATGGAGTGTTCACGACTTCCAGTAACTTTTCAAGGGATCAGCCGTG 1219
DB 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGATGTGAGAGGTGTTCCTTGGTCCATATGCCACAGGAT 1279
DB 341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCAACTATCAATGGGTGCCCTTATCAAGAGAGAGTCCCTATTCACGGCCAGAACT 1339
DB 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCGCCAGCAAAACATTTGGTGGTTTGACNCTCAAGAGGACCTTCCTGATGATGTATA 1399
DB 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAAGAAGTCAATCCAGCCATGATGACCAATCCAGTGTTCCTATGAACAATCCCCA 1459
DB 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACGGATGTAATTCATATTTACAAATTCGTAATTCGTAGACCGAGTGGAT 1519
DB 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAGATGCACAGTATGATGTTATCGAACAGATGTTGGACCGTTCCTTAAA 1579
DB 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAAGGTTCTCTCGAAGAAATG 1639
DB 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluMet 480
QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTCCAAATGGAGCTTCCACTAAGCAGCAACAA 1699
DB 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
QY 1700 CTATATATTGTTCAACGGCTGGGGTTCGCCAGCTCCCTTTTACACCGGTGTGATATTTAC 1759
DB 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGCGTGTGCTGAGTGTTCCTCCCGCAGACCCCTTACTGTGCTTGGATGGTCT 1819
DB 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCCACTGCAAGAGACGACACAGCAGACAGATATAGAAT 1879
DB 541 AlaCysSerArgTyrPheProThrAlaLysArgThrArgGlnAspIleArgAsn 560

```



```
QY 1880 GGAGACCCACTGACTGCTTCTGAGCTTACACCATGATAATCACCATGGCCACAGCCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisGlyHisSerPro 580
QY 1940 GAAGAGAGAAATCATCTATGCTGTAGAGAAATAGTAGCACATTTTGGAAATGCGAGTCCGAAG 1999
Db 581 GluGluArgGlyLeuLeuThrHisGlyValGluAsnSerSerThrPheLeuGluCysSerProlys 600
QY 2000 TCGCAGAGAGCGCTGCTTATTGGCAATTCAGAGCGGAAATCAAGAGCGAAAGAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTGGATCATCATATCATCAGCAGACAGATCAAGGCCCTTCTGCTAGCTAGTCTACAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGGATTACGGCAATTAACCTTCGCCATGCGGTGGAAATGCGGTTCATACAAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheLeuGlnThrLeu 660
QY 2180 CTTAAGGTAAACCTGGAAGTCAATTCAGCAGAGCAATTTGGAAGAACTTCTTCATAAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLysAsp 680
QY 2240 GATGATGGAGATGGCTCTAAGACCAAGAAATGTCATATAGATGACACCTAGCCAGAAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGTGACAGAGACTTCATGAGCTCATCAACACCCCAATCTCAACACAGATGATGAG 2359
Db 701 ValTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAAAGTTTGGAAAGGACCGAAACAAACGTCGGCAAGGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGAACAGTAACAAATGGAGACATTCACAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGlnAsnLysGlyArgAsnArgArg 760
QY 2480 ACCCAGCAATTTGAGAGGCGCCAGGAGGTGTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 2
US-10-262-538-10
; Sequence 10; Application US/10262538
; Publication No. US2003011324A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo et al
; TITLE OF INVENTION: NEUROFILIN/VEGFR-C/VEGFR-3 MATERIALS AND METHODS
; FILE REFERENCE: 28967/37564
; CURRENT APPLICATION NUMBER: US/10/262,538
; CURRENT FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-538-10

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-262-538-10 (1-771)
QY 200 ATGGCGCTGGTTAACTAGGATTCGTCTTTCTGGGGAGTATTACTTACAGCAAGACGA 259
```

```
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGAATGGGAAGAACAAATGTGCCAAGGCTGAAATTTATCTCAACAAAGAATGTTG 319
Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAACAATGTGATCACTTTCAATGGCTTGCCCAACAGCTCCAGTTATCATACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
QY 380 CTTTGTGATGAGAACCGAGTAGGCTGTATCTTGGAGCAAAAGATCAATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAATATATACAGGATTTTCAAAGATTTGTGGCCAGTATCTTACACCAAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATGAATGCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTCTAAATTTTCATCAAGGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATATACAGACTCATTGTACCCCTGTGGAAACGGGGCTTTTCATCCAAAT 619
Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTGAAATTTGGACATCATCTCAGAGACAAATATTTTAAAGCTGGAGAACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisIleProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAAAACGGCCGTGGGAAGAGTCCATATACCCCTAAGCTGCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGAGAAATATATCTCTGGAACCTCCAGCTGATTTTATGGGCGAGACTTTGCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCCGAACCTTTGGGCACCCACCACCAATCAGGACAGAGCAGAGTATTTCCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisIleHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCAATGATCCAAAGTTTCATTAGTGGCCACCTCATCTCAGAGAGTGAACAATCTCGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuLeuSerGluSerAspAsnProGluAsp 240
QY 920 GACAAGATATCTTTTCTTCCGTGAAATATGCAATAGATGGAGAACACTCTGGAAGAGCT 979
Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCACGCTAGAATAGTTCAGATATGCAAGAATCACTTTGGAGGGCCACAGAGTCTGGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGCAACAATTCCTCAAGCTCGTCTGATTTTGTCTCAGTCCAGGTCCTCAAAATGGC 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGTATGAATGTCAGGATGTATCTTAATGAACCTTTAAAGATCTTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTGTGTATGAGTGTGTTTACCACTTCCAGTAACTTTTCAAGGGATCAGCCGTG 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGAGTGTATGTGAGAAAGGTGTTCTCTGCTCCATATGCCACAGGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCCAACTATCAATCGGTGCTTATCAAGGAAGAGTCCCTTATCCACGCCCAAGGAAC 1339
```

Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCAGCAAAACATTTGGTGGTTTGGACTCTACAAAGACCTTCCTGATGATTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspValIle 400
QY 1400 ACCTTTGCAAGAAGTCATCCAGCCATGTAATCCAGTGTTCCTATGAACAATCGCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTCATCAAAACGGATGTAATTAATCAATTTACACAATTTGCTAGACCGAGTGGAT 1519
Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAGATGGACAGATGATGTTATGTTATTCGGAACAGATGTTGGGACCGTCTTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAGAGTTCCTGCTGGAAGAAATG 1639
Db 461 ValValSerIleProLysGluThrTyrAspLeuGluGluValLeuLeuGluGluMet 480
QY 1640 ACATGTTTTCGGGAACCGACTGCTATTTACGAATGGAGCTTTCCACTAAGCAGCAACAA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATTGTTCAAGCGCTGGGTTGCCAGTCCCTTTACACGGGTGCTAATTTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGCGTGTCTCAGTGTTCCTCGCCGAGACCTTACTGTCTGGTGGATGGTCT 1819
Db 521 GlyIleAlaCysAlaGluCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCCACTGCAAGAGACGCAAGACGCAAGATATAAGAAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTGTTCCAGACTTACACCATGATTAATCACCATGGCCACAGCCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisIleGlyHisSerPro 580
QY 1940 GAAGAGAGATCATCTATCGTGTAGAGATAGTAGCACATTTTGGAAATGCAGTCCGAG 1999
Db 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGCGCTGCTCTATTGGCAATTCAGAGCGCAATCAAGAGCGAAAGAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTCGATCATATCATCAGACAGATCAAGGCTTCTGCTAGTGTCTACAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGGATTCAGGCAATTAACCTCTCCATCGGTGGAACTGGTGTTCATACAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAAACCTGGAAGTCATTGACACAGAGCAATTTGGAAGAACTTCTTCATAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuHisLysAsp 680
QY 2240 GATGATGGAGATGGCTCTAAGACCAAGAAATGTCMAATAGCATGACACCTAGCCAGAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGTGTACAGAGATTCATGATGAGCTCATCAACCCCACTCTCAACACGATGATGAG 2359
Db 701 ValTyrTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAGATTTGGAAAGGGACCGAAACCAACGTCGCAAGGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740

QY 2420 CCAGGACACTAACAAATGAAGCACTTCAAGAAATAAGAAAGTAGAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCCACCAGTAATTTGAGAGGGCACCCAGGAGTGTCTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 3

US-09-946-374-310
; Sequence 310, Application US/09946374
; Publication No. US20030073129A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2830P1C1
; CURRENT APPLICATION NUMBER: US/09/946,374
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/098716
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098723
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098749
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098750
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098803
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098821
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098843
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/099536
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099596
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099598
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099602
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099642
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099741
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099754
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099763
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099792
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099808

;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099812
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099815
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099816
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/100385
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100388
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100390
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100584
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100627
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100661
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100662
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100664
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100683
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/100684
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/100710
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/100711
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/100848
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/100849
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/100919
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/100930
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/101014
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/101068
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/101071
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/101279
;; PRIOR FILING DATE: 1998-09-22
;; PRIOR APPLICATION NUMBER: 60/101471
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101472
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101474
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101475
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101476
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101477
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101479
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101738
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101741
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101743
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101915
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101916
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/102207
;; PRIOR FILING DATE: 1998-09-29

;; PRIOR APPLICATION NUMBER: 60/102240
;; PRIOR FILING DATE: 1998-09-29
;; PRIOR APPLICATION NUMBER: 60/102307
;; PRIOR FILING DATE: 1998-09-29
;; PRIOR APPLICATION NUMBER: 60/102330
;; PRIOR FILING DATE: 1998-09-29
;; PRIOR APPLICATION NUMBER: 60/102331
;; PRIOR FILING DATE: 1998-09-29
;; PRIOR APPLICATION NUMBER: 60/102484
;; PRIOR FILING DATE: 1998-09-30
;; PRIOR APPLICATION NUMBER: 60/102487
;; PRIOR FILING DATE: 1998-09-30
;; PRIOR APPLICATION NUMBER: 60/102570
;; PRIOR FILING DATE: 1998-09-30
;; PRIOR APPLICATION NUMBER: 60/102571
;; PRIOR FILING DATE: 1998-09-30
;; PRIOR APPLICATION NUMBER: 60/102684
;; PRIOR FILING DATE: 1998-10-01
;; PRIOR APPLICATION NUMBER: 60/102687
;; PRIOR FILING DATE: 1998-10-01
;; PRIOR APPLICATION NUMBER: 60/102965
;; PRIOR FILING DATE: 1998-10-02
;; PRIOR APPLICATION NUMBER: 60/103258
;; PRIOR FILING DATE: 1998-10-06
;; PRIOR APPLICATION NUMBER: 60/103314
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103315
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103328
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103395
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103396
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103401
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103449
;; PRIOR FILING DATE: 1998-10-06
;; PRIOR APPLICATION NUMBER: 60/103633
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/103678
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/103679
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/103711
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/104257
;; PRIOR FILING DATE: 1998-10-14
;; PRIOR APPLICATION NUMBER: 60/104987
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105000
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105002
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105104
;; PRIOR FILING DATE: 1998-10-21
;; PRIOR APPLICATION NUMBER: 60/105169
;; PRIOR FILING DATE: 1998-10-22
;; PRIOR APPLICATION NUMBER: 60/105266
;; PRIOR FILING DATE: 1998-10-22
;; PRIOR APPLICATION NUMBER: 60/105693
;; PRIOR FILING DATE: 1998-10-26
;; PRIOR APPLICATION NUMBER: 60/105694
;; PRIOR FILING DATE: 1998-10-26
;; PRIOR APPLICATION NUMBER: 60/105807

Alignment Scores:

Pred. No.:	1,05e-199	Length:	777
Score:	2245.50	Matches:	413
Percent Similarity:	73.67%	Conservative:	141
Best Local Similarity:	54.92%	Mismatches:	177
Query Match:	45.86%	Indels:	21
DB:	11	Gaps:	9

US-09-774-490-1	(1-2709)	x	US-09-946-374-310	(1-7777)
Qy	275	AAGAACAAATGTCGAAGCGTGAATAATTATCTCAACAAAGAATGTGGAAATCCACAATGTG	334	
Dd	39	LysGlnAsnIleProArgLeuLysLeuThrTyrllysAspLeuSerAsnSerCys	58	
Qy	335	ATCACTTTCAATGGCTTGCCCAACACGCTCCAGTTATCATACCTCTCTTTGGATGAGAA	394	
Dd	59	IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuLeuAspGlu	78	
Qy	395	CGGAGTAGGCTGTATGTGAGCAACAGGATCACATATTTTCATCGACTGCTTAATATC	454	
Dd	79	ArgGlyArgLeuLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu	98	
Qy	455	---AAGGATTTCAAAGATGTGTGCCCATGATCTTACACAGAGAGAGATGAATGCAAG	511	
Dd	99	AsnLysAsnPhelYsIleTyTrpProAlaAlaLysGluArgValGluLeuCysLys	118	
Qy	512	TGGCTCGAAAAGACATCCTGAAAGATGTGCTAATTTTCATCAAGGTACTTTAAGGCATAT	571	
Dd	119	LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPhelIleArgValLeuGlnProTyrl	138	
Qy	572	AATCAGACTCACATTGACGCTGTGGAACGGGGCTTTTCATCCAAATTGACCTTACATT	631	
Dd	139	AsnLysThrHisIleTyValCysGlyThrGlyAlaPheHisProlleCysGlyTyrlle	158	
Qy	632	GAAATTGGACATCATCTCGAGGACAATATTTTTAAGCTGGGAACTCACATTTGAAAAAC	691	
Dd	159	AspLeuGlyValTyrllysGluAspIlellePheLysLeuAspThrHisAsnLeuGluSer	178	
Qy	692	GGCGTGGGAAGAGTCCATATGACCCTAGCTCTGACAGCATCCCTTTTAATAGATGA	751	
Dd	179	GlyArgLeuLysCysProPheAspProGlnInProPheAlaSerValMetThrAspGlu	198	
Qy	752	GAATTAATCTCGAATCGAGCTGATTTTATGGCGGAGACTTGTCTATCTTCGGAAT	811	
Dd	199	TyrLeuTyrlserGlyThrAlaSerAspPheLeuGlyAspThrAlaPheThrArgSer	218	
Qy	812	CTTGGG-----CACCAACCCCNAATCAGACAGACGATGATTCAGGTGGCTC	862	
Dd	219	LeuglyProthrHisAspHisIeTyrlleArgThrAspIleSerGluHisTyrlTrpLeu	238	
Qy	863	AATGATCCAAGTTTATTAGTCCACCTCATCTCAGAGAGTGACATCTTCAAGATGAC	922	
Dd	239	AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrlasnProAspAsp	258	
Qy	923	AAAGTATATCTTTCTTCCTGAAATGCAATAGATGGAGAACACTCTGMAAAGCTACT	982	
Dd	259	LysIleTyrlPhePheArgGluSerSerGlnGluGlySerThrSerAspLysThrlle	278	
Qy	983	CACGCTAGAATFAGTATGACAGATATGATTTGGAGGGGACAGAAAGTCTGTGTAAT	1042	
Dd	279	LeuSerArgValGlyArgValCysLysAsnAspValGlyGlnArgSerLeuIleAsn	298	
Qy	1043	AAATGGHCAACATCTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCCAATGGCAT	1102	
Dd	299	LysTrpThrThrPheLeuLysAlaArgLeuilleCysSerIleProGlySerAspGlyAla	318	
Qy	1103	GACACTCATTTTGATGAATCGAGGATGATTCCTTAATGAATCTTAAAGATCCTAAAAAT	1162	
Dd	319	AspThrTyrlPheAspGluLeuGlnAspIleTyrlleLeuLeuProthrArghspGluArgAsn	338	
Qy	1163	CCAGTTGTATAGGAGTGTATACGATTCAGTAAACATTTTCAAGGGATCAGCGCTGTGT	1222	
Dd	339	ProValValTyrlGlyValPheThrThrThrThrSerSerIlePheLysGlySerAlaValCys	358	
Qy	1223	ATGTATACGATGATGTGTGAGAGGGTCTTCTTGTCTCATATGCCCCACAGGATGA	1282	
Dd	359	ValTyrlSerMetAlaAspIleArgAlaValPheAsnGlyProTyrlAlaHisLysGluSer	378	
Qy	1283	CCCAACTATCAATGGGTGCCTTATCAAGGAAGAGTCCCTTATCCAGCGCCAGGAATCTGT	1342	

QY 2420 CCAGGGAACAGTACAAATGGAAGCACTTACAGAAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 748 -----ProlystipLySHisMetGlnGluMetLysLysArgaenArgAG 763
QY 2480 ACCCAC---GAATTTCAGAGGCGACCCAGGAGTGTC 2512
Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775

RESULT 4

US-10-052-586-348
; Sequence 348, Application US/10052586
; Publication No. US20020127584A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C1
; CURRENT APPLICATION NUMBER: US/10/052,586
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059266
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063120
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063121
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063486
; PRIOR FILING DATE: 1997-10-21
; PRIOR APPLICATION NUMBER: 60/063540
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063541
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063544
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063564
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063734
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063870
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066120
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066466
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066772
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069335
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069425
; PRIOR FILING DATE: 1997-12-12
; PRIOR APPLICATION NUMBER: 60/069870
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/068017
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086023
; PRIOR FILING DATE: 1998-05-18
; PRIOR APPLICATION NUMBER: 60/086392
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/086486
; PRIOR FILING DATE: 1998-05-22

Db 379 AlaAspHisArgTrpValGlnThrAspGlyArgGlleProTyrProArgProGlyThrCys 398
QY 1343 CCAGCAAAACATTGGTGT---TTTGACTCTACAAAGGACCTTCCTGATGATGTATA 1399
Db 399 ProSerLysThrTyrAspProLeuLeuLysSerThrArgAspPheProAspValIle 418
QY 1400 ACCTTTGCAAGAGTACCCACCCATGATCAATCCAGTGTTCCTATGACAAATGCCCA 1459
Db 419 SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro 438
QY 1460 ATAGTGATCAAAACCGGATTAATATCAATTTACACAATTTGCTAGACCGAGTGGAT 1519
Db 439 ThrPheLysArgIleasnValAspTyrArgLeuThrGlnIleValAspHisValIle 458
QY 1520 GCAGAGATGGACAGTATGATTTATCGGAACAGATGTTGGACCGTCTCTAAA 1579
Db 459 AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
QY 1580 CTAGTTTCATTCCTAAGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAAGAAATG 1639
Db 479 ValValSerIleSerLysGluLysTrp---AsnMetGluGluValValLeuGluGluLeu 497
QY 1640 ACAGTTTTCGGGAACCGACTCTATTTCAGCAATGGAGCTTCCCAAGCAGCAACA 1699
Db 498 GlnIlePheLysHisSerSerIleLeuAsnMetGluLeuSerLeuLysGlnGln 517
QY 1700 CTATATATTGGTTCACCGCTGGGTGCCAGCTCCCTTTACACCGGTGATATTTAC 1759
Db 518 LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
QY 1760 GGAAGAGTGTGCTGAGTGTGCTCGCCGAGACCTTACTGCTGCTGGATGGTCT 1819
Db 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAsn 557
QY 1820 GCATGTTCTCGCTATTTCCTCCACTCAAGAGACCCACAGACAGATATGAAT 1879
Db 558 AlaCysSerArgTyrAlaProThrSerLysArgAlaArgArgLysAspValLysTyr 577
QY 1880 GGAGACCCACTGACTCTGTTTACAGTTCACCATGATATCACCATGGCCGACAGCCCT 1939
Db 578 GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GlnThrAla 596
QY 1940 GNAGAGAGATCATCTATGCTGTAGAGATAGTACACATTTTGGATGAGTCCGAG 1999
Db 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616
QY 2000 TCGCAGAGAGCGCTGCTCTATTGGCAATTCAGAGCGCAATGAGAGCGCAAGAGAG 2059
Db 617 SerGlnGlnAlaThrIleLysTrpTyrIleGlnArgSerGlyAspGluHisArgGluGlu 636
QY 2060 ATCAGATGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAGTCTCAA 2119
Db 637 LeuLysProAspGluArgIleIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln 656
QY 2120 CAGAGGATTCGGCAATTTACTCTGCGATCGGTGGAAACATGGGTTCATACAACTCTT 2179
Db 657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
QY 2180 CTTAAGGTAAACCTGGAAGTCAATTCACACAGGATTTGGAGAACTTCTTCATAAAGAT 2239
Db 677 ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu 696
QY 2240 GATGATGGAGATGGCTCTAAGACCAAAAGAAATGTCCAATAGCATGACACCTAGCCAGAG 2299
Db 697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
QY 2300 GTCTGGTACAGACTTCATGAGCTCATCAACCAACCCCAATCTCAACACAGATGATGAG 2359
Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerProAsnPhe---SerLeuAspGln 729
QY 2360 TTCTGTGAACAGTTTGGAAAGGACCGAAACACACCTCGGCAAGCCAGGACATACC 2419
Db 730 TyrCysGluGlnMetIrrPheArgGluLysArgGlnArgAsnLysGlyGly----- 747

QY 2420 CCAGGAACAGTAACAATGAAGCACCTTACAAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 748 -----ProLysTrpLysHisMetGlnGluMetLysLysLysArgAsnArgArg 763
QY 2480 ACCCAC---GAATTTGAGAGGGCCACCCAGGAGTGTCT 2512
Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775
RESULT 5
US-10-174-590-348
; Sequence 348, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; PRIORITY FILING DATE: 2002-06-18
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-348
Alignment Scores:
Pred. No.: 1,05e-199 Length: 777
Score: 2245.50 Matches: 413
Percent Similarity: 73.67% Conservative: 141
Best Local Similarity: 54.92% Mismatches: 177
Query Match: 45.86% Indels: 21
DB: Gaps: 9
US-09-774-490-1 (1-2709) x US-10-174-590-348 (1-777)
QY 275 AAGAACAAATGTGCCAAGGCTGAAATTTATCTTACAAGAAATGTTGGAAATCAACAATGTG 334
Db 39 LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuSerAsnSerCys 58
QY 335 ATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCATCTCTCTTTTGGATGAGAA 394
Db 59 IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuLeuAspGluGlu 78
QY 395 CGAGTAGTGGCTGATGTTGGCAAGCAAGATCAATATTTTCATTCGACCTGTTTAATATC 454
Db 79 ArgGlyArgLeuLeuLeuGlyAlaLysAspHisIlePheLeuLeuLeuValAspLeu 98
QY 455 ---AAGATTTCAAAAGATTGTGTCGCCAGATCTTTACACCAAGAGATGAATGCAAG 511
Db 99 AsnLysAsnPheLysLysIleTyrTrpProAlaAlaLysGluArgValGluLeuCysLys 118
QY 512 TGGCTCGAAAGACATCTCGAAGAAATGTGCTTAATTTTCATCAAGGTACTTAAGGCATAT 571
Db 119 LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr 138
QY 572 AATCAGACTCATTGTACGCTGTGGAACGGGGCTTTTCATCCAAATTTCCACCTACATT 631
Db 139 AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle 158
QY 632 GAAATTGGACATCATCTCTAGGACAATATTTTTTAAGCTGGAGAACTCACATTTTGAAC 691

Db 159 AspLeuGlyValTyrLysGluAspIlePheLysLeuAspThrHisAsnLeuGluSer 178
QY 692 GCGGTGGGAAGAGTCATATGACCCCTAAGCTGCTGACAGATCCCTTTTAAATAGATGA 751
Db 179 GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu 198
QY 752 GAATATATCTCTGGAACCTGACCTGATTTATGGCGGAGACTTTGCTATCTTCCGAAC 811
Db 199 TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer 218
QY 812 CTTGGG-----CACCACCAACCAATCAGGACAGACGATGATCCAGGTGGCTC 862
Db 219 LeuGlyProThrHisAspHisIleAspHisIleAspHisIleAspHisIleAspHisIle 238
QY 863 AATGATCCAAATTCATTAAGTCCCACTCATCTCAGAGAGTGACAACTCCTGAAGATGAC 922
Db 239 AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrAsnProAspAsp 258
QY 923 AAGTATATCTTTCTTCCTCGTCAAAATGCAATAGATGGAGACACTCTGGAAGCTACT 982
Db 259 LysIleTyrPhePhePheArgGluSerSerGlnGluGlySerThrSerAspLysThrIle 278
QY 983 CACGCTAGATAGGTGAGATATGCAAGATGACTTTGGAGGACACAGAACTCTGCTGAAT 1042
Db 279 LeuSerArgValGlyArgValCysLysAsnAspValGlyGlnArgSerLeuIleAsn 298
QY 1043 AAATGACAAATTCCTCAAGCTCGTGTGATTGCTCAGTCCAGGTCCTCAATGGCATT 1102
Db 299 LysTrpThrThrPheLeuLysAlaArgLeuIleCysSerIleProGlySerAspGlyAla 318
QY 1103 GACACTCATTTGATGAACCTGAGATGTATTCCTTAATGAATTTAAAGATCCTAAAT 1162
Db 319 AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArgAsn 338
QY 1163 CCAGTTGTATGAGTGTATGACGACTCCAGTAAATTTCAAGGATCAGCCGTGTGT 1222
Db 339 ProValValTyrGlyValPheThrThrSerSerIlePheLysGlySerAlaValCys 358
QY 1223 ATGTATAGATGAGTGTAGAGAGGGTTCCTTGTGTCATATGCCACAGGATGA 1282
Db 359 ValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysGluSer 378
QY 1283 CCCAACTATCATGGTGCCTTATCAAGAGAGTCCCTATCCAGCCGACGACTGT 1342
Db 379 AlaAspHisArgTrpValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys 398
QY 1343 CCAGCAAAACATTTGGTGGT---TTTGACTCTCAAAAGGACCTTCCTGATGATGTATA 1399
Db 399 ProSerLysThrTyrAspProLeuLysSerThrArgAspPheProAspValIle 418
QY 1400 ACCTTTGCAAGATCATCCAGCATGTACAATCCAGTGTTCCTATGAACCAATCCGCA 1459
Db 419 SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyPro 438
QY 1460 ATAGTGATCAAAACGGATTAATATCAATTTACAAATTTGCTAGACCGATGGAT 1519
Db 439 ThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIle 458
QY 1520 GCAGAGATGAGCAGTATGATGTTTATCGGAACAGATGTTGGGACCCCTCTTAAA 1579
Db 459 AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
QY 1580 GTAGTTTCAATTCCTAAGAGACTTGGTATGATTTAGAGAGTTCTGCTGGAAGAAATG 1639
Db 479 ValValSerIleSerLysGluLysTrp---AsnMetGluGluValValLeuGluGluLeu 497
QY 1640 ACAGTTTTCGGAACCGACTCTATTTCCAGCAATGGACCTTCCTCAAGCAGCAACA 1699
Db 498 GlnIlePheLysHisSerSerIleIleLeuAsnMetGluLeuSerLysGlnGln 517
QY 1700 CTATATATGTTCAACGGCTGGGTGGCCAGCTCCCTTTACACGGGTGTATTTAC 1759

Db 518 LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
QY 1760 GGGAAAGCGTGTCTGAGTGTGGCTCGCCGAGACCCCTTACTGTCTGCTGGATGGTCT 1819
Db 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAsn 557
QY 1820 GCATGTTCTCTGCTATTTTCCCACTGCAAGAGAGCGCACAGACGACAAAGATATAAAGAT 1879
Db 558 AlaCysSerArgTyrAlaProThrSerLysArgAlaAlaArgGlnAspValLysTyr 577
QY 1880 GGAGACCCACTGACTCTACTGTTTCAGACTTACACCATGATATACCATGGCCACAGCCCT 1939
Db 578 GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GluThrAla 596
QY 1940 GAAGAGAGATCATCTATGTTGATAGATAGTACACATTTTGGATGTCAGTCCGAG 1999
Db 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616
QY 2000 TCGCAGAGAGCGTGTCTATTGGCAATTCAGAGGCGAAATGAAGACGCAAAAGAGAG 2059
Db 617 SerGlnGlnAlaThrIleLysTrpTyrIleGlnArgSerGlyAspGluHisArgGluGlu 636
QY 2060 ATCAGATGGATGATCATATCATCAGGACAGATCAAGCCCTTCTGCTAGTAGTCTCAA 2119
Db 637 LeuLysProAspGluArgIleIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln 656
QY 2120 CAGAAGATTCCAGCAATTACCTCTGCATCGGTGGACACATGGTTCATACAACTCTT 2179
Db 657 LysLysAspSerGlyMetTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
QY 2180 CTTAAGGTAAACCCCTGGAGTCAATTCAGACACAGCATTTGGAGAACTTCTTCATAAAGAT 2239
Db 677 ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu 696
QY 2240 GATGATGAGATGCTCTTAAGACCAAAAGATGTCATAGCATCACACCTAGCCAGAG 2299
Db 697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
QY 2300 GTCTGGTACAGAGACTTCATGCAGCTCATCAACACCCCATCTCAACACCATGGATGAG 2359
Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln 729
QY 2360 TTCTGTGAACAAGTTTGGAAAGGACCGCAAAACAACTCGGCAAGCCAGGACATACC 2419
Db 730 TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly----- 747
QY 2420 CCAGGAAACAGTAAACAATGGAACCTTCAAGAAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 748 -----ProLysTrpLysHisMetGlnMetLysLysLysArgAsnArgArg 763
QY 2480 ACCCAC---GAATTTGAGAGGCGACCCAGGAGTGTCT 2512
Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775

RESULT 6

US-10-176-758-348

; Sequence 348, Application US/10176758

; Publication No. US20030008353A1

GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; TITLE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3430R1C104

657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrile 676
2180 CTTAGGTAAACCTGGAAGTCTATTGACACAGAGCATTTGGAGAACTCTTTCATTAAGAT 2239
677 ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu 696
2240 GATGATGGAGATGGCTCTAAGACCAAGAAAGATGTCCAATAGCATGACACCTAGCCAGAG 2299
697 HisGluGluGluGlnValLys-----AspLeuLeuAlaGluSerArg 710
2300 GTCTGTACAGACATTCATGAGCTCATCAACACCCCAATCTCAACACGATGATGAG 2359
711 LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln 729
2360 TTCTGTGAACAAGTTGGAAAGGACCGAAACAAACACGTCGCAAGGCGCAGGACATACC 2419
730 TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly----- 747
2420 CCAGGGAACAGTAACAATGGAAGCACTTACAGAAATAAAGAAAGGTAGAAACAGGAGG 2479
748 -----ProLysTrpLysHisMetGlnGluMetLysLysLysArgAsnArgArg 763
2480 ACCAC---GAATTTGAGAGGCGACCCAGGAGTGC 2512
764 HisHisArgAspLeuAspGluLeuProA-GAlaVal 775

RESULT 7

US-10-175-737-348
; Sequence 348, Application US/10175737
; Publication No. US20030013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C50
; CURRENT APPLICATION NUMBER: US/10/175,737
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-737-348

Alignment Scores:
Pred. No.: 1,05e-199 Length: 777
Score: 2245.50 Matches: 413
Percent Similarity: 73.67% Conservative: 141
Best Local Similarity: 54.92% Mismatches: 177
Query Match: 45.86% Indels: 21
DB: 15 Gaps: 9

US-09-774-490-1 (1-2709) x US-10-175-737-348 (1-777)

275 AAGAACAAATGTGCAAGGCTGAAATTTATCTCAAAAGAAATGTGGAATCCAAATGTG 334
39 LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuSerAsnSerCys 58
335 ATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCATACCTTCTTTTGGATGGAA 394
59 IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuLeuAspGluGlu 78

395 CGAGTAGGCTGTATGTTGGACAAAGATCACATATTTTCATTCGACTCGTGAATATC 454
79 ArgGlyArgLeuLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
455 ---AAGATTTTCAAAGATTGTGTGCCAGTATCTTACACAGAGAGATGAATGCAAG 511
99 AsnLysAsnPheLysIleTyrTrpProAlaAlaLysGluArgValGluLeuCysLys 118
512 TGGCTGGAAGACATCTCTGAAAGATGTCTTAATTTTCATCAAGGTACTTTAAGCCATAT 571
119 LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr 138
572 AATCAGACTCACTTGTAGCCTGTGGAACGGGGCTTTTCATCCAAATTTGACCTACATT 631
139 AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle 158
632 GAATTTGACATCATCTCTGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTTGAANAAC 691
159 AspLeuGlyValTyrLysGluAspIlePheLysLeuAspThrHisAsnLeuGluSer 178
692 GGCCTGGAGAGTCCATATGACCTAAGCTGTCAGACGATCCCTTTTAAATAGATGA 751
179 GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu 198
752 GAATTTACTCTGGAACCTGACACTGATTTTATGGGCGAGACTTTTGTCTATCTTCCGAAT 811
199 TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer 218
812 CTTGGG-----CACCACCACCATCAGGACAGCAGCATGATTCAGGTGGCTC 862
219 LeuGlyProThrHisAspHisIleTyrIleArgThrAspIleSerGluHisTyrTrpLeu 238
863 AATGATCCAAAGTTTCATTAGTCCACCTCATCTCAGAGAGTGACAACTCTGGAAGATGAC 922
239 AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrAsnProAspAsp 258
923 AAAGTATATCTTTTCTCCGTGAAATGCAATAGATGGAGAACTCTGGAAGACTACT 982
259 LysIleTyrPhePhePheArgGluSerSerGlnGluGlySerThrSerAspLysThrIle 278
983 CAGCTAGATAGTACAGATATGCAAGATACATCTTGGAGGCGCAGAGTCTGCTGAT 1042
279 LeuSerArgValGlyArgValCysLysAsnAspValGlyGlyGlnArgSerLeuLeuAsn 298
1043 AAATGGACAACATCTCTCAAGCTCGTCTGATTTTGTCTCAGTCCAGGTCGCAAAATGGCAT 1102
299 LysTrpThrThrPheLeuLysAlaArgLeuLeuCysSerIleProGlySerAspGlyAla 318
1103 GACACTCATTTTGTAGTAATCGAGATGTATTCTTAATGAATTTTAAAGATCTTAAAT 1162
319 AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArg 338
1163 CCAGTTGTATATGAGTGTTCAGCTTCAGTACATCTTTCAGGGATCAGCCGTGTGT 1222
339 ProValValTyrGlyValPheThrThrSerSerIlePheLysGlySerAlaValCys 358
1223 ATGTATAGCATGATGATGTCGAGAGGCTGTTCTTGTGTCATATGCCACAGAGATGA 1282
359 ValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysGluSer 378
1283 CCCAACTATCAATGGTGCCTTTATCAAGGAAGAGTCCCTATCCACGCCAGGAACTGT 1342
379 AlaAspHisArgTrpValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys 398
1343 CCCAGCAAAACATTTGGTGGT---TTTACTCTCAAAAGGACCTTCTCGATGATCTATA 1399
399 ProSerLysThrTyrAspProLeuIleLysSerThrArgAspPheProAspAspValIle 418
1400 ACCTTTGCAAGAAGTCTCCAGCCCATGTACAATCCAGTGTTCCTTATCATCAACAATGCCCA 1459
419 SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyPro 438
1460 ATAGTGTATCAAAACGGATGTAAATTTACAAATTTACAAATTTGTCTAGACCGAGTGGAT 1519

```

Db 439 ThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIle 458
QY 1520 GCAGAGATGGACAGTATCATGTTATGTTTATCGGAACAGATGTTGGGACCCGCTCTTAAA 1579
Db 459 AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
QY 1580 GTAGTTTCAATTCCTAAGAGACTGCTGATGATTTAGAGAGGTTCTGCTGGAAGAAATG 1639
Db 479 ValValSerIleSerLysGluLysTrp---AsnMetGluGluValValLeuGluGluLeu 497
QY 1640 ACAGTTTTTCGGAACCGACTCTATTTCAGCAATGGAGCTTCCACTAAGCAGCAACA 1699
Db 498 GlnIlePheLysHisSerIleLeuAsnMetGluLeuSerLeuLysGlnGln 517
QY 1700 CTATATATTGTTCAACGGCTGGGTTGCCAGCTCCCTTTACACCGGTGCTATTTAT 1759
Db 518 LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
QY 1760 GGAAGAGCTGTGCTGAGTGTGCTGCTGCCGAGACCTTACTGTGTTGGGATGTTCT 1819
Db 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAsn 557
QY 1820 GCATGTTCTCGCTATTTTCCACTGCAAGACAGCAGCAGCAGCAGCAGATATAAAT 1879
Db 558 AlaCysSerArgTyrAlaProThrSerLysArgArgAlaArgGlnAspValLysTyr 577
QY 1880 GGAGACCCACTGACTCAGTCTTACAGCTTACACCATGATATACCATCGGCCACAGCCCT 1939
Db 578 GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GluThrAla 596
QY 1940 GAAGAGAGATCATCTATGTTAGAGATAGTAGACATTTTTCGAATTCAGTCCGAG 1999
Db 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616
QY 2000 TCGCAGAGCGCTGCTTATGCGCAATTCAGAGCGCAATTCAGAGCGCAATTCAGAGCGCA 2059
Db 617 SerGlnGlnAlaThrIleLysTrpTyrIleGlnArgSerGlyAspGluHisArgGluGlu 636
QY 2060 ATCAGAGTGGATGATCATATCATCAGACAGATCAAGGCTTCTCTAGTGTCTACAA 2119
Db 637 LeuLysProAspGluArgIleIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln 656
QY 2120 CAGAAGGATTCAGCAATACCTCTGCCATCGGTGGGACATCGGTTTCATACAACTCTT 2179
Db 657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
QY 2180 CTTAGGTAACTCGAAGTCACTTACACAGACGATTTGGAGAACTTCTTCATTAAGAT 2239
Db 677 ValLysLeuThrLeuAsnValIleGluAsnGlnMetGluAsnThrGlnArgAlaGlu 696
QY 2240 GATGATGGAGTGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAGAG 2299
Db 697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
QY 2300 GTCTGCTACAGACTTCATGCTCAGCTCATCAACCCCACTCTCAACACGATGGATGAG 2359
Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln 729
QY 2360 TTCTGTGAACAGTTTGGAAAGGACCGAACAACCTCGGCAAGGCCCGCAGGACATACC 2419
Db 730 TyrCysGluGlnMetTrpHisArgGluLysArgArgGlnArgAsnLysGlyGly----- 747
QY 2420 CCAGGAACAGTAAACAAATGGAAGCACATTACAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 748 -----ProLysTrpLysHisMetGlnGlnMetLysLysLysArgAsnArgArg 763
QY 2480 ACCAC---GAATTTGAGAGGCGCACCAGGAGTGC 2512
Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775

```

RESULT 8

US-10-173-706-348

```

; Sequence 348, Application US/10173706
; Publication No. US20030022293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173,706
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-173-706-348

```

```

Alignment Scores:
Pred. No.: 1,05e-199 Length: 777
Score: 2245.50 Matches: 413
Percent Similarity: 73.67% Conservative: 141
Best Local Similarity: 54.92% Mismatches: 177
Query Match: 45.86% Indels: 21
DB: 15 Gaps: 9

```

US-09-774-490-1 (1-2709) x US-10-173-706-348 (1-777)

```

QY 275 AAGAACATGTGCCAAGGCTGAATATTATCTTACAAAGAAATGTTGGAATCCAAACAATGTG 334
Db 39 LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuLeuSerAsnSerCys 58
QY 335 ATCACTTTCAATGCTGTGGCCACAGCTCCAGTTCATCATCTCTCTTTGGATGAGAA 394
Db 59 IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuLeuAspGluGlu 78
QY 395 CGGAGTAGCTGTATGTTGGAGCAAGCATCATATATTTTCATTCGACCTGTTAAATC 454
Db 79 ArgGlyArgLeuLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
QY 455 ---AAGGATTTTCAAAAGATTTGTGTGGCCAGTATCTTACACAGAGAGATGAATGCAAG 511
Db 99 AsnLysAsnPheLysLysIleTyrTrpProAlaAlaLysGluArgValGluLysCysLys 118
QY 512 TGGCTCGAAAGACATCTCGAAAGATGTGCTAATTTTCATCAAGGTACTTTAAGGCATAT 571
Db 119 LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr 138
QY 572 AATCAGACTCACTGTAGCTGTGGACCGGGCTTTTCATCCCAATTTGACCTACATTT 631
Db 139 AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle 158
QY 632 GAAATTCGACATCATCTCGAGGACAAATATTTTAAAGCTGGAGAACTCATTGTAANAAC 691
Db 159 AspLeuGlyValTyrLysGluAspIleIlePheLysLeuAspThrHisAsnLeuGluSer 178
QY 692 GGCGTGGGAAGAGTCCATATGACCCCTTAAGCTGCTGACAGCATCCCTTTTAAATAGATGA 751
Db 179 GlyArgLeuLysCysProPheAspProGlnProPheAlaSerValMetThrAspGlu 198
QY 752 GAATTATCTCTGGAAGTGGAGCTGATTTTATGCGGCGAGACTTGTGCTATCTTCGGAAT 811
Db 199 TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer 218

```

812 CTTGGG-----CACCACCAATCAGGACAGACGATGATTCCAGGTGGCTC 862
Db |||||
219 LeuGlyProThrHisAspHisHisThrIleArgThrAspIleSerGluHisThrLeu 238
Qy |||||
863 AATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGACAATCTCTGAAGATGAC 922
Db |||||
239 AsnGlyAlaLeuPheIleGlyThrPhePheIleProAspThrAsnProAspAsp 258
Qy |||||
923 AAGATATATCTTTCTTCCTCGTGAATGCAATAGATGAGAGACACTCTGGAAGAACTACT 982
Db |||||
259 LysIleThrPhePheArgGluSerGlnGluSerThrSerAspLysThrIle 278
Qy |||||
983 CACGCTAGATAGTTCAGATATGCAAGATGCTTTGGAGGACACAGAGTCTGTGTAAT 1042
Db |||||
279 LeuSerArgValGlyArgValCysLysAsnAspValGlyGlyLysSerLeuLeuLeu 298
Qy |||||
1043 AATGGACAACTTCTCAAGCTCGTCTGATTGCTCAGTCCAGGTCCTCAATGGCAAT 1102
Db |||||
299 LysThrThrPheLeuLysAlaArgLeuIleCysSerIleProGlySerAspGlyAla 318
Qy |||||
1103 GACACTCATTTGATGAAGTCAAGTGTATTCCTTAATGAATCTTAAGATCTCTAAAT 1162
Db |||||
319 AspThrThrPheAspGluLeuGlnAspIleThrLeuLeuProThrArgAspGluArg 338
Qy |||||
1163 CAGATTGTATATGGAGTGTTCAGCTTCCAGTAACTTTCAAGGATCAGCGGTGT 1222
Db |||||
339 ProValValThrGlyValPheThrThrSerSerIlePheGlySerAlaValCys 358
Qy |||||
1223 ATGTATAGATGATGATGAGAGGTGTTCTTGTGTCCTATGTCCTCCACAGGATGGA 1282
Db |||||
359 ValThrSerMetAlaAspIleArgAlaValPheAsnGlyProThrAlaHisLysGluSer 378
Qy |||||
1283 CCAACTATCAATGGTGCTTATCAAGAGAGTCCCTATCCAGGCGCAGAACTGT 1342
Db |||||
379 AlaAspHisArgThrValGlnThrAspGlyArgGlyProThrArgProGlyThrCys 398
Qy |||||
1343 CCAGCAAACTATTTGGTGT---TTTGACTCTCAAAAGGACTTCTCTGATGATGTTATA 1399
Db |||||
399 ProSerLysThrThrAspProLeuIleLysSerThrArgAspPheProAspValIle 418
Qy |||||
1400 ACCTTTGCAAGATGATCCAGCATGTATCAATCCAGTGTTCCTATGCAATCCGCCA 1459
Db |||||
419 SerPheIleLysArgHisSerValMetThrLysSerValThrProValAlaGlyGlyPro 438
Qy |||||
1460 ATAGTATCAAAAGGATGATTAATCAATTAATCAATTAATCAATTAATCAATTAATCA 1519
Db |||||
439 ThrPheLysArgIleAsnValAspThrArgLeuThrGlnIleValValAspHisValIle 458
Qy |||||
1520 CGAAGATGGACAGTATGATGTTATGTTATCGGAACAGATGTTGGGACCGTCTTAA 1579
Db |||||
459 AlaGluAspGlyGlnThrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
Qy |||||
1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAAGAGTTTCTGCTGGAAGAAATG 1639
Db |||||
479 ValValSerIleSerLysGluLysThr---AsnMetGluValValLeuGluLeu 497
Qy |||||
1640 ACAGTTTTCGGAACCGACTGCTATTTCAAGCATGAGACTTTCACCTAAGCAGCAGCAA 1699
Db |||||
498 GlnIlePheLysHisSerIleLeuAsnMetGluLeuSerLeuLysGlnGln 517
Qy |||||
1700 CTATATATGTTTCAACGCTGGGTTCGCCAGTCCCTTTTACCCGGTGTATTTTAC 1759
Db |||||
518 LeuThrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrThr 537
Qy |||||
1760 GGAAGACGCTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTTTGGATGCTTCT 1819
Db |||||
538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProThrCysAlaThrAspGlyAsn 557
Qy |||||
1820 GATGTTCTCGCTATTTTCCACTGCAAGAGACGACAGCAGCAGCAAGATATAAGAAAT 1879
Db |||||
558 AlaCysSerArgThrAlaProThrSerLysArgAlaArgGlnAspValLysThr 577
Qy |||||
1880 GGAGACCCACTGACTCTGCTTCCAGACTTACACCATGATAATCACCATGCCACAGCCCT 1939

578 GlyAspProIleThrGlnCysThrAspIleGluAspSerIleSerHis---GluThrAla 596
Qy |||||
1940 GAAAGAGAAATCATCTATGTTGTTAGATAGTACACATTTTGGATGCGAGTCCGAG 1999
Db |||||
597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616
Qy |||||
2000 TCGCAGAGAGCGTGTCTATTGGCAATTCAGAGGCGCAATTAAGAGCGCAAAAGAGAG 2059
Db |||||
617 SerGlnGlnAlaThrIleLysThrPheGlnArgSerGlyAspGluHisArgGluGlu 636
Qy |||||
2060 ATCAGAGTGGATGATCATCATCATCAGGACAGACATCAAGCCCTTCTGCTAGTGTACAA 2119
Db |||||
637 LeuLysProAspGluArgIleIleLysThrGluThrGlyLeuLeuIleArgSerLeuGln 656
Qy |||||
2120 CAGAAGATTTCAGCAATTACTCTGCTATGCTGCGGTGGACATGGTTCATACAACTCTT 2179
Db |||||
657 LysLysAspSerGlyMetThrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
Qy |||||
2180 CTTAAGGTAAACCTGGAAGTCAATGACACAGACATTTGGAAGAACTTCTTCATAAAGAT 2239
Db |||||
677 ValLysLeuThrLeuAsnValIleGluAsnGlnMetGluAsnThrGlnArgAlaGlu 696
Qy |||||
2240 GATGATGAGATGCTCTAAGACCAACCAAGAAATGTCATGATGACACCTAGCCAGAG 2299
Db |||||
697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
Qy |||||
2300 GTCTGTTACAGAGACTTCTGCTGCTCATCAACCAACCCCAATCTCAACAGTGTAGAG 2359
Db |||||
711 LeuArgThrLysAspThrIleGlnIleLeuSerProAsnPhe---SerLeuAspGln 729
Qy |||||
2360 TTCTGTGCAAGTTTGGAAAGGACCGCAAAACCAACGTCGCAAGGCGCAGGACATACC 2419
Db |||||
730 TyrCysGluGlnMetThrPheArgGluLysArgGlnArgAsnLysGlyGly----- 747
Qy |||||
2420 CCAGGCAACATAACAATGGAAGCACTTACAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db |||||
748 -----ProLysThrLysHisMetGlnGluMetLysLysLysArgAsnArgArg 763
Qy |||||
2480 ACCAC---GAATTTGAGAGGCGCCAGGAGTGTCT 2512
Db |||||
764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775

RESULT 9

US-10-175-738-348

; Sequence 348, Application US/10175738

; Publication No. US2003002294A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C45

; CURRENT APPLICATION NUMBER: US/10/175,738

; PRIORITY FILING DATE: 2002-06-19

; Prior application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 348

; LENGTH: 777

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-175-738-348

Alignment Scores:

Pred. No.:	1,05e-199	Length:	777
Score:	2245.50	Matches:	413
Percent Similarity:	73.67%	Conservative:	141
Best Local Similarity:	54.92%	Mismatches:	177
Query Match:	45.86%	Indels:	21
DB:	15	Gaps:	9
US-09-774-490-1 (1-2709) x US-10-175-738-348 (1-777)			
QY	275	AAGAACAAATGTCAGGCTGAAATATCTCTACAAAGAAATGTTGGAATCCAAATGTC	334
DB	39	LYSGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuSerAsnSerCys	58
QY	335	ATCACTTTCAATGGCTGGCCCAACAGCTCCAGTTATCATACCTTCTCTTTGGATGAGAA	394
DB	59	IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuLeuAspGlu	78
QY	395	CSGAGTAGGCTGTGTTGGACCAAGGATCACATATTTTCATTCGACTGGTTAATATC	454
DB	79	ArgGlyArgLeuLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu	98
QY	455	---AAGGATTTTCAAGATTTGTGGCCAGTATCTTACACAGAGAGATGAATGCAAG	511
DB	99	AsnLysAsnPheLysLysIleTyrTrpProAlaAlaLysGluArgValGluLeuCysLys	118
QY	512	TGGCTGGAAAGACATCTCGAAGATGTGCTAATTTTCATCAAGGTACTTAAGCATAT	571
DB	119	LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr	138
QY	572	AATCAGACTCACTTACGCTGTGGAAGCGGGCTTTTCATCCAAATTTGCACCTACAT	631
DB	139	AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle	158
QY	632	GAAATGGACATCATCTCGAGACAAATATTTTAAAGCTGGGAACTCAATTTTCAAAAC	691
DB	159	AspLeuGlyValTyrLysGluAspIlePheLysLeuAspThrHisAsnLeuGluSer	178
QY	692	GGCCTGGAGAGTCCATATGACCTTAAGCTGTGACAGCATCCCTTTTAAATAGATGA	751
DB	179	GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu	198
QY	752	GAAATTAATCTGGAACCTGACCTGATTTTATGGGCGAGACTTGTCTATCTCCGAAT	811
DB	199	TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer	218
QY	812	CTTGGG-----CACCAACCAATCAGACAGACAGCAGATGATTCAGGTGGCTC	862
DB	219	LeuGlyProThrHisAspHisTyrIleArgThrAspIleSerGluHisTyrTrpLeu	238
QY	863	AATGATCCAAAGTTCAATAGTGCCCACTCTCAGAGAGTGACAATCCTGGAAGATGAC	922
DB	239	AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrAsnProAspAsp	258
QY	923	AAAGTATATCTTTCTTCGGTGAAATGCAATAGATGAGAACACTCTGGAAAGCTACT	982
DB	259	LysIleTyrPhePhePheArgGluSerSerGlnGluGlySerThrSerAspLysThrIle	278
QY	983	CACGTAGATAGGTTCAGATATGCAAGATGACTTTGGAGGCGACAGAACTCTGTGTAAT	1042
DB	279	LeuSerArgValGlyArgValCysLysAsnAspValGlyGlnArgSerLeuIleAsn	298
QY	1043	AAATGGACAACTTCTCAAGCTCGTCTGATTTGCTCAGTGCCAGGTCCAAATGGCAT	1102
DB	299	LysTrpThrThrPheLeuLysAlaArgLeuIleCysSerIleProGlySerAspGlyAla	318
QY	1103	GACACTCATTTTGTGAAGTGTATGTTTCTTAATGACTTTTAAAGATCTTAAAT	1162
DB	319	AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArgAsn	338
QY	1163	CCAGTTGTATGGAGTGTTCAGACTTCCAGTACATTTTCAAGGGATCAGCGTGTGT	1222
DB	339	ProValValTyrGlyValPheThrThrSerSerIlePheLysGlySerAlaValCys	358

QY	1223	ATGTATAGCATGAGTGTGATGTGAGAGGTGTCTTGGTCCATATGCCCACAGGATGA	1282
DB	359	ValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysGluSer	378
QY	1283	CCCAACTATCAATGGGTGCTTATCAAGGAGAGTCCCTATCCACGCCCGAGAACTTGT	1342
DB	379	AlaAspHisArgTrpValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys	398
QY	1343	CCCGACAAAACATTTGGTGGT---TTTGACTCTCAAGAGACCTTCTCTGATGATGTATA	1399
DB	399	ProSerLysThrTyrAspProLeuLysSerThrArgAspPheProAspValIle	418
QY	1400	ACCTTTGCAAGAGTCTATCCAGCCATGACATCCAGTTTCTCTATGAACAATCCGCA	1459
DB	419	SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro	438
QY	1460	ATAGTGATCAAAACGGATGTAATATCAATTTACACAAATTTGCTAGACCGAGTGGAT	1519
DB	439	ThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIle	458
QY	1520	GCAGAGATGCACAGTATGATGTTTATCGAAGACAGATGTTGGACCGCTTCTTAAA	1579
DB	459	AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys	478
QY	1580	GTAATTTCAATTCCTAAGAGACTTGTGATGATTTAGAGAGTTTCTCTCGAAGAAATG	1639
DB	479	ValValSerIleSerLysGluLysTrp---AsnMetGluGluValValLeuGluLeu	497
QY	1640	ACAGTTTTCGGGAAACCGACTGCTATTTTCAGCAATGGAGCTTCCACTAAGCAGCAAA	1699
DB	498	GlnIlePheLysHisSerSerIleIleLeuAsnMetGluLeuSerLeuLysGlnGlnGln	517
QY	1700	CTATATATTTGTTTCAACGGCTGGGTTGCCAGCTCCCTTTTACACCGGTGTGATTATAC	1759
DB	518	LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr	537
QY	1760	GGGAAAGCGTGTGCTGAGTGTGCTCGCCCGAGACCTTACTGTCTGGATGGTTCT	1819
DB	538	GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAsn	557
QY	1820	GCATGTTCTCGTATTTTCCACCTGCAAGAGAGCGACAGACGACAGATATAGAAT	1879
DB	558	AlaCysSerArgTyrAlaProThrSerLysArgAlaArgArgGlnAspValLysTyr	577
QY	1880	GGACACCACTGACTCTACTGTTACACTTACCATGATATCACCATGGCCACAGCCCT	1939
DB	578	GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GluThrAla	596
QY	1940	GAAGAGAAATCATCTATGTTGTAGAGAAATAGTAGACATTTTGGATGAGTCCGAG	1999
DB	597	AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys	616
QY	2000	TCGACAGAGCGGTGCTTATTTGGCAATTCAGAGGCGAAATGAAGCGGAAAGAGAG	2059
DB	617	SerGlnGlnAlaThrIleLysTrpTyrIleGlnArgSerGlyAspGluHisArgGlu	636
QY	2060	ATCAGGTGATGATCATATCATCAGACAGATCAAGCGCTTCTGCTACGTAGTCTCAA	2119
DB	637	LeuLysProAspGluArgIleIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln	656
QY	2120	CAGAGGATTCAGCAATTAACCTCTGCTGCGGTGGAAATGCGGTTCATACAAACTCT	2179
DB	657	LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle	676
QY	2180	CTTAAGTAACTTGGAGTCAATTCATGACAGCATTTTGAAGAACTTCTTCTATAAGAT	2239
DB	677	ValLysLeuThrLeuAsnValIleGluAsnGlnMetGluAsnThrGlnArgAlaGlu	696
QY	2240	GATCATGGAGTGGCTCTAAGACCCAAAGAAATGTCCAATAGCATGACACCTAGCCAG	2299
DB	697	HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg	710
QY	2300	GTCTGGTACAGACTTTCATGCAGCTCATCAACCCCAATCTCTCAACAGCATGGATGAG	2359

Qy	1640	ACAGTTTTTCGGGAACCGACGTGCTATTATTTACGCAATGAGCGCTTTCCACTAAGACGACACAA	1699
Db	498	GlnIlePheIysHisSerSeriIleLeuAsnMetGluLeuSerLeuIysGlnGlnGln	517
Qy	1700	CTATATATTGCTTCAACGGCTGGGGTGGCCAGCTCCCTTTTACACGGTGTGATATTATTC	1759
Db	518	LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr	537
Qy	1760	GGGAACCGCTGTGCTGAGTGTGCTCGCCGAGACCCCTTACTGTGCTTCGGATGGTTC	1819
Db	538	GlyIysAlaCysAlaAspCysLeuAlaArgAspProTyrCysAlaTpaAspGlyAsn	557
Qy	1820	GCATGTTCTCGCTATTTTCCACTGTCGAAGAGACGACAAAGACGACAAAGATATAAGAAAT	1879
Db	558	AlaCysSerArgTyrAlaProThrSerIysArgAlaArgGlnAspValIysTyr	577
Qy	1880	GGAGACCCACTGACTACTGTTCCAGCTTACACCATGATTAATCACCATTGCCACACGCCCT	1939
Db	578	GlyAspProIleThrGlnCysTpaSpIleGluAspSerIleSerHis---GluThrAla	596
Qy	1940	GAAGACGAGATCATCTATGTGTAGAGAAATAGTAGACATTTTGAATCGACTCCGAAG	1999
Db	597	AspGluIysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProIys	616
Qy	2000	TCGCAGAGACGCTGTCTATTGGCAATCCAGAGCGCAATGAAGACGCGAAAGAGAG	2059
Db	617	SerGlnGlnAlaThrIleLysTptTyrIleGlnArgSerGlyAspGluHisArgGluGlu	636
Qy	2060	ATCAGAGTGATGATCATATCATCAGACAGATCAAGGCCCTTCTGCTAGCTGTACAA	2119
Db	637	LeuIysProAspGluArgIleIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln	656
Qy	2120	CAGAAGGATTCAGCAATTACCTCTGCATCGCGTGGAAACATGGTTCATACAACTCTT	2179
Db	657	LysIysAspSerGlyMetTyrTyrCysIysAlaGlnGluHisThrPheIleHisThrIle	676
Qy	2180	CTTAAGGTAAACCTCGGAAGTCATTGACACAGACGATTTGGAAGAACTTCTTCATAAGAT	2239
Db	677	ValIysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu	696
Qy	2240	GATGATGGAGATGGCTCTAAGACCAAGAATAATGTCCAATAGCATGACACTGACCGAAG	2299
Db	697	HisGluGluGlyGlnValIys-----AspLeuLeuAlaGluSerArg	710
Qy	2300	GTCTGTGTACAGACITTCATCGACTCATCAACCCCAATCTCAACAGCATGGATGAG	2359
Db	711	LeuArgTyrIysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln	729
Qy	2360	TTCTGTCAACAAAGTTTGGAAAGGGACCGCAAAACAACGTCCGCAAGGCCGAGACATACC	2419
Db	730	TyrCysGluGlnMetTrpHisArgGluIysArgArgGlnArgAsnIysGlyGly-----	747
Qy	2420	CCAGGGAACAGTAAACAATCGAAGCACTTACAGAATAATAGAAAGAGTGAACACGAGG	2479
Db	748	-----ProlysTtpIysHisMetGlnGluMetIysIysIysArgAsnArgArg	763
Qy	2480	ACCCAC---GAATTTGAGGGCACCCAGAGTGTCT	2512
Db	764	HisHisArgAspLeuAspGluLeuProArgAlaVal	775

RESULT 11

US-10-176-482-348

US 10-170-402-348
; Sequence 348, Application US/10176482

; Sequence 348, Application 03/10/2003
; Publication No. US20030022296A1

; GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.

APPLICANT: Chen, Jian

APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.

APPLICANT: Gurney, Austin L.

APPLICANT: Pan, James

APPLICANT: Smith, Victoria

```

; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C70
; CURRENT APPLICATION NUMBER: US/10/176,482
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-482-348

```

Alignment Scores:		
Pred. No.:	1,059-199	777
Score:	2245.50	413
Percent Similarity:	73.67%	141
Best Local Similarity:	74.92%	177
Query Match:	45,86%	21
DB:	15	9
		Gaps:
		Length:
		Matches:
		Conservative:
		Mismatches:
		Indels:

US-09-774-490-1 (1-2709) x US-10-176-482-348 (1-777)

	Qy	275	AAGAACAAATGTGCCAAGGCTGAATAATTATCCTACAAAGAATAATGTTGGAAATCCCAACAATGTG	334
	Db	39	LysGlnAsnIleProArgLeuLysLeuThrTyrIysAspLeuLeuSerAsnSerCys	58
	Qy	335	ATCACITTCATGTGCTTGCCCAAAGCTCCAGTTATCATCTTCCTTTTGCATGAGGAA	394
	Db	59	IleProPheLeuGlySerSerGluClyLeuAspPheGlnThrLeuLeuAspGluGlu	78
	Qy	395	CGGAGTAGGCTGTATGTGGAGCAAAAGGATCACATATTTTCATTCGCAGCTGGTAAATATC	454
	Db	79	ArgGlyArgLeuLeuLeuGlyAlaIysAspHisIlePheLeuLeuSerLeuValaspLeu	98
	Qy	455	-- --AAGGATTTTCAAAGATTGTGTGGCCAGTATCTTTACACAGAGAAGATGAATGCAAG	511
	Db	99	AsnLysAsnPheLysLysIleTy-TripProAlaLysGluArgValGluLeuCysLys	118
	Qy	512	TGGGCTGGAAAAGACATCCTGAAAGAATGTGCTTAATTTTCATCAAGGTACTTTAAGGCATAT	571
	Db	119	LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr	138
	Qy	572	AATCAGACTCACTTGATCGCTGTGGACGGGGGCTTTTCATCCAATTTGCACCTTACATT	631
	Db	139	AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle	158
	Qy	632	GAAATTGGACATCATCTCGAGGACAAATATTTTTAAAGCTGGAGAACTTCATATTTGAAAC	691
	Db	159	AspLeuGlyValTyrLysGluAspIleIlePheLysLeuAspThrHisAsnLeuGluSer	178
	Qy	692	GGCCGTGGGAAGAGTCCATATGACCCTAAAGCTGTGCAGCAGATCCCTTTTAAATAGATGA	751
	Db	179	GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu	198
	Qy	752	GAATTATACTCTGAACTGCGAGCTGATTTTATGGGGCGAGACTTTTGCTATCTTCGGAAC	811
	Db	199	TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer	218
	Qy	812	CTTGGG-----CACCACCACCAATCAGGACAGCAGCATGATTCAGGTGGGCTC	862
	Db	219	LeuGlyProThrHisAspHisHisTyrIleargThrAspIleSerGluHisTyrTrpLeu	238
	Qy	863	AATGATCCAAAGTTTCATTAGTGGCCCACTCATCTCAGAGAGTGACAATCCTCGAAAGAT	922
	Db	239	AsnGlyAlaLysPheIleGlyThrPhePheIleProaspThrTyrAsnProaspAsp	258
	Qy	923	AAAGTATACITTTTCTTCCTGGAAAAATGCAATAGATGAGAACACTCTGSAAGAACTACT	982
	Db	259	LvsIleTvrPhePhePheAcgGluSerSerGlnGluGlvSerThrSerAspLysThrIle	278

Db 39 LysGlnAsnIleProArgLeuIleLysLeuThrTyrLysAspLeuLeuLeuSerAsnSerCys 58
QY 335 ATCACTTTCAATGGCTGGCCACAGCTCCAGTTATCATACATCTCTCTTTGGATGAGGAA 394
Db 59 IleProPheLeuGlySerSerGluGlyLeuAspPheGlnThrLeuLeuLeuAspGluGlu 78
QY 395 CGGAGTAGCTGTATGTTGGAGCAAGAGATCATATTTTCATTTGACCTGTTAATATC 454
Db 79 ArgGlyArgLeuLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
QY 455 ---AAGGATTTCAAAGATGTGTGGCCAGTATCTTACACAGAGAGATGAATGCAAG 511
Db 99 AsnLysAsnPheLysIleTyrTrpProAlaAlaLysGluArgValGluLeuCysLys 118
QY 512 TGGCTGGAAGACATCTCGAAGAAATGTGTAATTTTCATCAAGTACTTAAGGCATAT 571
Db 119 LeuAlaGlyLysAspAlaAsnThrGluCysAlaAsnPheIleArgValLeuGlnProTyr 138
QY 572 AATCAGACTCACTTGACCTGTGGAACGGGGCTTTTCATCCAAATTTGCACCTACATT 631
Db 139 AsnLysThrHisIleTyrValCysGlyThrGlyAlaPheHisProIleCysGlyTyrIle 158
QY 632 GAAATTTGACATCTCTCAGACAAATATTTTAAAGCTGGAGAACTCAATTTGAAAC 691
Db 159 AspLeuGlyValTyrLysGluAspIlePheLysLeuAspThrHisAsnLeuGluSer 178
QY 692 GCCCTGGAAGAGTCCATATGACCTCAACCTAGCTGCTGACAGCATCCCTTTTATAGATGA 751
Db 179 GlyArgLeuLysCysProPheAspProGlnGlnProPheAlaSerValMetThrAspGlu 198
QY 752 GAATATATCTCTGGAACCTGACTGATTTTATGGGGAGAGCTTTGCTATCTTCCGAAT 811
Db 199 TyrLeuTyrSerGlyThrAlaSerAspPheLeuGlyLysAspThrAlaPheThrArgSer 218
QY 812 CTGGG-----CACCAACCCCAATCAGACAGACGACCATGATTCAGGTGGCTC 862
Db 219 LeuGlyProThrHisAspHisIleTyrIleArgThrAspIleSerGluHisTyrTrpLeu 238
QY 863 AATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGACAACTCTCAAGATGAC 922
Db 239 AsnGlyAlaLysPheIleGlyThrPhePheIleProAspThrTyrAsnProAspAsp 258
QY 923 AAGTATATCTTTCTTCGTGAAATGCAATAGATGGAGAACACTCTGGGAAACCTACT 982
Db 259 LysIleTyrPhePheArgGluSerSerGlnGluGlySerThrSerAspLysThrIle 278
QY 983 CACGCTAGATAGTGCAGATGCAAGATGACTTTGGAGGCAAGAGTCTGTGAT 1042
Db 279 LeuSerArgValGlyArgValCysLysAsnAspValGlyGlyGlnArgSerLeuIleAsn 298
QY 1043 AAATGGACAACATCTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCCAAATGGCAT 1102
Db 299 LysTrpThrThrPheLeuLysAlaArgLeuIleCysSerIleProGlySerAspGlyAla 318
QY 1103 GACATCAATTTGATGAAGTCAAGATGATATTCCTTAATGAATTTAAAGATCTTAAAT 1162
Db 319 AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArgAsn 338
QY 1163 CCAGTTGTATAGAGTGTTCAGACTTCCAGTACATTTTCAAGGGATCAGCGTGTGT 1222
Db 339 ProValValTyrGlyValPheThrThrThrSerSerIlePheLysGlySerAlaValCys 358
QY 1223 ATGTATAGCATGATGTGAGAGGGTGTTCCTTTGTTGCTATATGCTCCACAGGATGGA 1282
Db 359 ValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysGluSer 378
QY 1283 CCAACTATCAATGGTGGCTTATCAAGAGAGTCCCTATCCACGCCAGGAACTTGT 1342
Db 379 AlaAspHisArgTrpValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys 398
QY 1343 CCCAGCAACAACTTTGGTGGT---TTTGACTCTACAAAGGACCTTCTCTGATGATGTTATA 1399
Db 399 ProSerLysThrTyrAspProLeuIleLysSerThrArgAspPheProAspAspValIle 418

QY 1400 ACCTTTGCAAGAAGTCATCCAGCCATGTAACAATCCAGTGTTCCTTATCAACAATCGCCA 1459
Db 419 SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro 438
QY 1460 ATAGTGATCAAAACGGATGTAATATCAATTTTACAAATTTGCTGTAGACCGATGGAT 1519
Db 439 ThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIle 458
QY 1520 GCAGAAGATGCACAGTATGATGTTATCGGACAGATGTTGGACCGTCTCTTAAA 1579
Db 459 AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
QY 1580 GTAGTTTCAATTCCTAAGGACACTTGGTATGATATTAGAAAGAGTTCTCTCGAAGAAATG 1639
Db 479 ValValSerIleSerLysGluLysTrp---AsnMetGluGluValValLeuGluGluLeu 497
QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTCACCAATGAGAGCTTTCACCTAAGCAGCAACA 1699
Db 498 GlnIlePheLysHisSerSerIleIleLeuAsnMetGluLeuSerLeuLysGlnGlnGln 517
QY 1700 CTATATATTGTTCAACGGCTGGGTGCCAGCTCCCTTTTACACCGGTGTGATATTAC 1759
Db 518 LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
QY 1760 GCGAAAGCGTGTGCTGAGTGTGCTCGCCCGAGACCTTACTGTGCTTGGATGGTGTCT 1819
Db 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAsn 557
QY 1820 GCATGTTCTCCTATTTTCCACTGCAAGAGACGACACAGCAGCAAGATATAGAAT 1879
Db 558 AlaCysSerArgTyrAlaProThrSerLysArgAlaArgGlnAspValLysTyr 577
QY 1880 GGAGACCACTGACTACTGTTACAGCTTACACCATGATATCACCATGGCCACAGCCT 1939
Db 578 GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis---GluThrAla 596
QY 1940 GAAGAGAATCATCTATGTGTAGAGAAATAGTAGCACATTTTGGATGAGTCCGGAAG 1999
Db 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616
QY 2000 TCGCAGAGAGCGCTGCTTATTTGGCAATTCAGAGGCGAAATGAAGACGCAAGAGAG 2059
Db 617 SerGlnGlnAlaThrIleLysTrpTyrIleGlnArgSerGlyAspGluHisArgGluGlu 636
QY 2060 ATCAGATGGATGATCATATCATCAGACAGATCAAGSCCTTCTGCTACGTAGTCTACA 2119
Db 637 LeuLysProAspGluArgIleIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln 656
QY 2120 CAGAAGGATTCAGGCAATTAACCTCTGCCATGCGGTGGAACATGGGTTCATACAACTCTT 2179
Db 657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
QY 2180 CTTAAGTAACCTCGGAAGTCATTGACACAGACATTTGGAAGAACTTCTTCTATAAGAT 2239
Db 677 ValLysLeuThrLeuAsnValIleGluAsnGlnMetGluAsnThrGlnArgAlaGlu 696
QY 2240 GATCATGGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAGAG 2299
Db 697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
QY 2300 GTCTGGTACAGACTTTCATCGAGCTCATCAACCCACCTCATCTCAACAGCATGGATGAG 2359
Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln 729
QY 2360 TTCTGTGAACAGTTTCGAAAGGACCGAAACACAGCTCGGCAAGCCAGGACATACC 2419
Db 730 TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly----- 747
QY 2420 CCAGGGAACAGTAACAATCGAAGCACTTACAGAAATTAAGAAAGTAGAAGACAGAGG 2479
Db 748 -----ProLysTrpLysHisMetGlnMetLysLysLysArgAsnArg 763

Db 319 AspThrTyrPheAspGluLeuGlnAspIleTyrLeuLeuProThrArgAspGluArgAsn 338
QY 1163 CAGTTGTATATGAGAGTGTTCAGACTCCAGTAACATTTTCAAGGGATCAGCGTGTGT 1222
Db 339 ProValValTyrGlyValPheThrThrSerSerIlePheLeuGlySerAlaValCys 358
QY 1223 ATGTATACATGAGTGTGAGAGGGTGTTCCTTGGTCCATATGCCACAGGGATGGA 1282
Db 359 ValTyrSerMetAlaAspIleArgAlaValPheAsnGlyProTyrAlaHisLysSer 378
QY 1283 CCCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTATCCACGGCCAGCAACTGT 1342
Db 379 AlaAspHisArgTyrValGlnTyrAspGlyArgIleProTyrProArgProGlyThrCys 398
QY 1343 CCCAGCAAAACATTTGGTGGT---TTTGACTCTACAAGAGGACCTTCCTGATGATTTATA 1399
Db 399 ProSerLysThrTyrAspProLeuIleLysSerThrArgAspPheProAspValIle 418
QY 1400 ACCTTTGCAAGAGTCCAGCCCATGACCAATCCAGTGTTCCTATGAACAATCGCCA 1459
Db 419 SerPheIleLysArgHisSerValMetTyrLysSerValTyrProValAlaGlyGlyPro 438
QY 1460 ATAGTATCAAAACCGGATTAATTAATTAACAATTTACAAATTTGCTAGACCGAGTGGAT 1519
Db 439 ThrPheLysArgIleAsnValAspTyrArgLeuThrGlnIleValValAspHisValIle 458
QY 1520 GCAAGATGAGCAGTATGATTTATTTATCGGAACAGATGTTGGGACCGTCTTTAAA 1579
Db 459 AlaGluAspGlyGlnTyrAspValMetPheLeuGlyThrAspIleGlyThrValLeuLys 478
QY 1580 GTAGTTTCAATTCCTAAGAGAGTCTGGTATGATTTAGAGAGGTCTGCTGGAAGAAATG 1639
Db 479 ValValSerIleSerLysGlyLysTyrP---AsnMetGluGluValValLeuGluLeu 497
QY 1640 ACAGTTTTCGGAACCGACTGCTATTTCCAGCAATGGAGCTTCCACTAAGCAGCAACAA 1699
Db 498 GlnIlePheLysHisSerSerIleLeuLeuAsnMetGluLeuSerLeuLysGlnGln 517
QY 1700 CTATATATTTGTTCAACGGCTGGGTGCGGCTCCAGTCCCTTTACACGGTGTGATTTAC 1759
Db 518 LeuTyrIleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
QY 1760 GGAAGAGCGTGTGAGTGTTCCTCGCCGAGACCTTACTGCTTGGATGCTTCT 1819
Db 538 GlyLysAlaCysAlaAspCysLysLeuAlaArgAspProTyrCysAlaThrAspGlyAsn 557
QY 1820 GCATGTTCTCGCTATTTTCCCACTGCAAGAGACGCACAGACGACGAAGATATAAGAAAT 1879
Db 558 AlaCysSerArgTyrAlaProThrSerLysArgAlaArgArgGlnAspValLysTyr 577
QY 1880 GGAGACCCACTACTGCTTACATGTTCCAGCTTACACCATGATATACCATGCGCCACAGCCCT 1939
Db 578 GlyAspProIleThrGlnCysTyrAspIleGluAspSerIleSerHis---GluThrAla 596
QY 1940 GAAGAGAGAAATCATCTATGCTGTAGAGATAGTAGCACATTTTTCGAATGCACTCCGAG 1999
Db 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616
QY 2000 TCCGAGAGCGCTGCTGTATTGGCAATTCAGAGCGCGAAATGAAGAGCGAAAGAGAG 2059
Db 617 SerGlnGlnAlaThrIleLysTyrPtyrIleGlnArgSerGlyAspGluHisArgGluGlu 636
QY 2060 ATCAGATGGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGCTAGTCTACAA 2119
Db 637 LeuLysProAspGluArgIleIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln 656
QY 2120 CAGAAGGATTTCAGGCAATTAACCTCTGCGCATGCGTGGAAACATGGGTTCATACAACTCTT 2179
Db 657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
QY 2180 CTTAAGGTAAACCTGGAAGTCAATGACACAGAGCAATTTGGAAAGAACTTCTTCAATAAGAT 2239
Db 677 ValLysLeuThrLeuAsnValIleGluAsnGlnMetGluAsnThrGlnArgAlaGlu 696

QY 2240 GATGATGAGATGGCTCTTAAGACCAAGAAATGTCTAATAGATGACACCTAGCCAGAAG 2299
Db 697 HisGluGluGlyGlnValLys-----AspLeuLeuAlaGluSerArg 710
QY 2300 GTCTGTACAGAGACTTCATCAGCTCATCAACACCCCAATCTCAACACGATGGATGAG 2359
Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerProAsnPhe---SerLeuAspGln 729
QY 2360 TTCTGTGAACAAGTTTGGAAAAGGACCGAAACCAACGTCGCGCAAGCCGAGGACATACC 2419
Db 730 TyrCysGluGlnMetTyrHisArgGluLysArgGlnArgAsnLysGlyGly----- 747
QY 2420 CCAGGACACAGTAAACAATGGAACACCTTACAAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 748 -----ProLysTyrIleHisMetGlnGluMetLysLysLysArgAsnArgArg 763
QY 2480 ACCCAC---GAATTTGAGAGGCGCACCCAGGAGTGTCT 2512
Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775

RESULT 15

US-10-180-557-348
; Sequence 348, Application US/10180557
; Publication No. US20030022301A1

GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian
; APPLICANT: Deanoyers, Luc

; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C147

; CURRENT APPLICATION NUMBER: US/10/180.557
; CURRENT FILING DATE: 2002-06-25

; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 348
; LENGTH: 777

; TYPE: PRT
; ORGANISM: Homo Sapien

US-10-180-557-348

Alignment Scores:
Pred. No.: 1,05e-199 Length: 777
Score: 2245.50 Matches: 413
Percent Similarity: 73.67% Conservative: 141
Best Local Similarity: 54.92% Mismatches: 177
Query Match: 45.86% Indels: 21
Gaps: 9

US-09-774-490-1 (1-2709) x US-10-180-557-348 (1-777)

QY 275 AAGAACAAATGTCACAGGCTGAAATTTCTTACAAAGAAATGTTGGAATCCAAACAATGTG 334
Db 39 LysGlnAsnIleProArgLeuLysLeuThrTyrLysAspLeuLeuSerAsnSerCys 58
QY 335 ATCACTTTCAATGCTTGGCCCAACAGCTCCAGTTCATCATCTTCTTTGATGAGGAA 394
Db 59 IleProPheLeuGlySerSerGluGlyLeuPheGlnThrLeuLeuAspGluGlu 78
QY 395 CGGAGTAGGCTGTATGTTGGAGCAAGGATCACATATTTTCATTGACCTGTTAATATC 454
Db 79 ArgGlyArgLeuLeuGlyAlaLysAspHisIlePheLeuLeuSerLeuValAspLeu 98
QY 455 ---AAGGATTTTCAAAAGATTGTGTGGCCAGTATCTTACACGAGAGAGATGAATGCAAG 511

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 09:54:26 ; Search time 417 Seconds
(without alignments)
11308.390 Million cell updates/sec

Title: US-09-774-490-1
Perfect score: 4896
Sequence: 1 aatctttttttatcatgatg.....agggtttttttcttaataacc 2709

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Xgapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 5580241 seqs, 870357830 residues

Total number of hits satisfying chosen parameters: 11160482

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p.model -DEV=xlh
-O=/cgn2_1/USPTO spoil/US09774490/runat_03082003 095417 25381/app query fasta_1.2887
-DB=Pending_Patents AA_Main -QFMT=fastan -SUFFIX=rapm -MINMATCH=0.1 -LOOPCL=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=DIOSUM62 -TRANS=human40.cdi
-LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09774490 @CGN 1 1 628 @runat_03082003 095417 25381 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Pending Patents AA Main:
1: /cgn2_6/ptodata/2/paa/PCTUS COMB.pcp.*
2: /cgn2_6/ptodata/2/paa/US06 COMB.pcp.*
3: /cgn2_6/ptodata/2/paa/US07 COMB.pcp.*
4: /cgn2_6/ptodata/2/paa/US080 COMB.pcp.*
5: /cgn2_6/ptodata/2/paa/US081 COMB.pcp.*
6: /cgn2_6/ptodata/2/paa/US082 COMB.pcp.*
7: /cgn2_6/ptodata/2/paa/US083 COMB.pcp.*
8: /cgn2_6/ptodata/2/paa/US084 COMB.pcp.*
9: /cgn2_6/ptodata/2/paa/US085 COMB.pcp.*
10: /cgn2_6/ptodata/2/paa/US086 COMB.pcp.*
11: /cgn2_6/ptodata/2/paa/US087 COMB.pcp.*
12: /cgn2_6/ptodata/2/paa/US088 COMB.pcp.*
13: /cgn2_6/ptodata/2/paa/US089 COMB.pcp.*
14: /cgn2_6/ptodata/2/paa/US090 COMB.pcp.*
15: /cgn2_6/ptodata/2/paa/US091 COMB.pcp.*
16: /cgn2_6/ptodata/2/paa/US092 COMB.pcp.*
17: /cgn2_6/ptodata/2/paa/US093 COMB.pcp.*
18: /cgn2_6/ptodata/2/paa/US094 COMB.pcp.*
19: /cgn2_6/ptodata/2/paa/US095 COMB.pcp.*
20: /cgn2_6/ptodata/2/paa/US096 COMB.pcp.*
21: /cgn2_6/ptodata/2/paa/US097A COMB.pcp.*
22: /cgn2_6/ptodata/2/paa/US097B COMB.pcp.*
23: /cgn2_6/ptodata/2/paa/US098 COMB.pcp.*
24: /cgn2_6/ptodata/2/paa/US099A COMB.pcp.*
25: /cgn2_6/ptodata/2/paa/US099B COMB.pcp.*
26: /cgn2_6/ptodata/2/paa/US100 COMB.pcp.*
27: /cgn2_6/ptodata/2/paa/US101 COMB.pcp.*
28: /cgn2_6/ptodata/2/paa/US102 COMB.pcp.*

29: /cgn2_6/ptodata/2/paa/US103 COMB.pcp.*
30: /cgn2_6/ptodata/2/paa/US104 COMB.pcp.*
31: /cgn2_6/ptodata/2/paa/US60 COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	4201	85.8	771	1	PCT-US00-41943-24	Sequence 24, Appl
2	4201	85.8	771	1	PCT-US02-07826-284	Sequence 284, App
3	4201	85.8	771	8	US-08-488-935-2	Sequence 2, Appli
4	4201	85.8	771	8	US-08-489-057-2	Sequence 2, Appli
5	4201	85.8	771	20	US-09-694-085-24	Sequence 24, Appl
6	4201	85.8	771	22	US-09-791-537-31106	Sequence 31106, A
7	4201	85.8	771	26	US-10-067-632-54	Sequence 54, Appl
8	4201	85.8	771	26	US-10-097-340-284	Sequence 284, App
9	4201	85.8	771	28	US-10-247-671-164	Sequence 164, App
10	4201	85.8	771	28	US-10-362-538-10	Sequence 10, Appl
11	4201	85.8	771	31	US-09-791-537-101922	Sequence 184, App
12	4057.5	82.9	772	22	US-09-791-537-101922	Sequence 101922,
13	4035.5	82.4	772	22	US-09-791-537-72548	Sequence 72548, A
14	3989.5	81.5	772	16	US-09-362-167-32	Sequence 32, Appl
15	3989.5	81.5	772	22	US-09-791-537-128280	Sequence 128280,
16	3784.5	77.3	772	22	US-09-791-537-111919	Sequence 111919,
17	3478.5	71.0	666	22	US-09-791-537-62757	Sequence 62757, A
18	3207.5	65.5	778	22	US-09-791-537-105918	Sequence 105918,
19	2245.5	45.9	777	19	US-09-582-419A-1	Sequence 1, Appli
20	2245.5	45.9	777	24	US-09-946-374-310	Sequence 310, App
21	2245.5	45.9	777	26	US-10-006-041A-310	Sequence 310, App
22	2245.5	45.9	777	26	US-10-006-063A-310	Sequence 310, App
23	2245.5	45.9	777	26	US-10-006-116A-310	Sequence 310, App
24	2245.5	45.9	777	26	US-10-006-117A-310	Sequence 310, App
25	2245.5	45.9	777	26	US-10-006-130A-310	Sequence 310, App
26	2245.5	45.9	777	26	US-10-006-172A-310	Sequence 310, App
27	2245.5	45.9	777	26	US-10-006-485A-310	Sequence 310, App
28	2245.5	45.9	777	26	US-10-006-746A-310	Sequence 310, App
29	2245.5	45.9	777	26	US-10-006-768A-310	Sequence 310, App
30	2245.5	45.9	777	26	US-10-006-818A-310	Sequence 310, App
31	2245.5	45.9	777	26	US-10-006-856A-310	Sequence 310, App
32	2245.5	45.9	777	26	US-10-007-194A-310	Sequence 310, App
33	2245.5	45.9	777	26	US-10-007-236A-310	Sequence 310, App
34	2245.5	45.9	777	26	US-10-011-671A-310	Sequence 310, App
35	2245.5	45.9	777	26	US-10-011-692A-310	Sequence 310, App
36	2245.5	45.9	777	26	US-10-011-795A-310	Sequence 310, App
37	2245.5	45.9	777	26	US-10-011-795B-310	Sequence 310, App
38	2245.5	45.9	777	26	US-10-011-833A-310	Sequence 310, App
39	2245.5	45.9	777	26	US-10-012-064A-310	Sequence 310, App
40	2245.5	45.9	777	26	US-10-012-101B-310	Sequence 310, App
41	2245.5	45.9	777	26	US-10-012-121A-310	Sequence 310, App
42	2245.5	45.9	777	26	US-10-012-137A-310	Sequence 310, App
43	2245.5	45.9	777	26	US-10-012-149A-310	Sequence 310, App
44	2245.5	45.9	777	26	US-10-012-231A-310	Sequence 310, App
45	2245.5	45.9	777	26	US-10-012-237A-310	Sequence 310, App

ALIGNMENTS

RESULT 1
PCT-US00-41943-24
; GENERAL INFORMATION:
; Sequence 24, Application PC/TUS0041943
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: NOVEL SEMAPHORIN DOMAINS
; FILE REFERENCE: 00786-441W01
; CURRENT APPLICATION NUMBER: PCT/US00/41943
; PRIOR FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/694,085
; PRIOR FILING DATE: 2000-10-20

;; PRIOR APPLICATION NUMBER: US 60/164,056
;; PRIOR FILING DATE: 1999-11-08
;; NUMBER OF SEQ ID NOS: 48
;; SOFTWARE: FASTSEQ for Windows Version 4.0
;; SEQ ID NO 24
;; LENGTH: 771
;; TYPE: PRT
;; ORGANISM: Homo sapiens
PCT-US00-41943-24

Alignment Scores:

Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US00-41943-24 (1-771)

QY 200 ATGGGCTGGTTAACTAGGATGCTGCTCTTTCTGGGAGTATTACTTACAGCAAGCA 259
DB 1 MetGlyTrpLeuThrArgileValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGAATGGGAAGAACAAATGTGCCAAGCTGAAATTTATCTACAAAGAAATGTTG 319
DB 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAATGTGATCACTTTCAATGGCTTGGCCAAAGCTCCAGTTATCATACCTTC 379
DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyrHisThrPhe 60
QY 380 CTTTGGATGAGAACGGTAGGCTGTATGTGGAGCAAGATACATATTTTCATTC 439
DB 61 LeuLeuAspGluLysArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGTTAAATATCAAGATTTTCAAAAGATTTGTGCCAGTATCTTACACCAAGAGA 499
DB 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATGAATGAATGGCTGGAAAGACATCTCTGAAGATGTCTTAATTTTCATCAAGTA 559
DB 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCATTGTACGCTGTGGACGGGGCTTTTCATCAATTT 619
DB 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTTGAAATTCGACATCATCTCGAGGACATATTTTAAAGCTGGAGAACTCA 679
DB 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAAACGGCGGTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGCATCCCTT 739
DB 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGGAATATTATCTCGACTGAGCTGATTTTATGGGCGAGACTTTGCT 799
DB 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTCCGAACCTCTGGGCACACACCACTTACAGACAGACAGCATGATTTCCAGGTGG 859
DB 201 IlePheArgThrLeuGlyHisHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCATATGATCCAAAGTTATAGTCCACCTCATCTCAGAGAGTGACAAATCCTGAAGAT 919
DB 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAGTATATCTTCTCGTGAATGCAATGATGAGACACTCTCGAAAGCT 979
DB 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGAAATAGGTGAGATATGCAAGATAGCTTTTGGAGGGGCACAGAGTCTGCTG 1039

DB 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGCAACAATCTCTCAAAAGCTGCTGATTGCTCAGTGCAGCCAGGTCCCAATGCC 1099
DB 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTACACACTATTGATGAATCAGATGATTTCTTAATGAATCTTTAAAGATCCTAAA 1159
DB 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGATATGGAGTGTTTACGACTTCCAGTAACATTTTCAAGGGATCAGCCGTG 1219
DB 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGATATAGCATGATGATGAGAAGGTGCTCTTGGTCCATATGCCACAGGGAT 1279
DB 341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCAACTATCAATGGTGCCTTTATCAAGGAAGAGTCCCTATCCACGCCAGGAACT 1339
DB 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCCAAGCAAAACATTTGGTGGTTTCACTCTCAAGAGGACCTTCTCTGATGATGTATA 1399
DB 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAAGAGTATCCAGCCATGTACAACTCCAGTGTCTTCTATGAAACAATGCCCA 1459
DB 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACGGATGTAATTTATCAATTTACAAATTTGCTGATAGCCGAGTGTAT 1519
DB 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAGATGACAGTATGATGTTATGTTTATCGGAACAGATGTTGGGACCGTCTTTAAA 1579
DB 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATCTCTAAGGAGACTTGGTATGATTTAGNAGAGGTTCTGCTGGAGAAATG 1639
DB 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTTCGGGAACCGACTGCTATTTTACCAATGAGAGCTTTCCACTAAGCAGCAACA 1699
DB 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
QY 1700 CTATATATTGTTCAACGGCTGGGTTGCCAGCTCCCTTTTACACCGGTGTGATTTTAC 1759
DB 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGGTGTGCTGAGTGTGGCTCGCCGAGACCTTACTGCTGTTGGATGGTGTCT 1819
DB 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTGCTATTTTCCCACTGCAAGAGAGCAGCAAGACCAAGATATAGAAT 1879
DB 541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGATCTGTTGAGCTTACCATCATATATCATCATGTCAGCCCT 1939
DB 561 GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisIleGlyHisSerPro 580
QY 1940 GAAAGAGATCATCTATGTTAGATAGTATAGTACATTTTGGATGTCAGTCCGAG 1999
DB 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGAGCGCTGGTCTATTGGCAATTCAGAGGCGAAATGAAGACGAAAAAGAGAG 2059
DB 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAGTCTACA 2119

Db 621 lleArgValAspAspHisIlelleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAGGATTACCAATACCTCTGCCATCGGTGGACATGGGTTTCATACAAACTCTT 2179
Db 641 GlnlysAspSerGlyAsnTyrLeuGlyHisIleValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAGGTAACCTCGAGTCAATCAGACAGAGCAATTTGGAAGACTTCTTCATAAAGAT 2239
Db 661 LeuIysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuHisIysAsp 680
QY 2240 CATGATGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATCACACCTAGCCAGAG 2299
Db 681 AspAspGlyAspGlySerIysThrIysGluMetSerAsnSerMetThrProSerGlnlys 700
QY 2300 GTCTGGTCAGAGACTTCATGAGCTCATCAACACCCCAATCTCAACACAGATGATCAG 2359
Db 701 ValTprTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGACAGATTGGAAAGGACCGAAACACCTCGGCAAGGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTrpIysArgAspArgIysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGAACAGTAACAAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnIysTrpIysHisIleuGlnGluAsnIysLysGlyArgAsnArgArg 760
QY 2480 ACCACGAAATTGAGAGGCCACCCAGGAGTGTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 2

PCT-US02-07826-284
; Sequence 284, Application PC/TUS0207826
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030PC
; CURRENT APPLICATION NUMBER: PCT/US02/07826
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 284
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-07826-284

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US02-07826-284 (1-771)

QY 200 ATGGCTGGTTAACTAGGATTGTCGTCTTTCTGGGGAGGTATTACTTACAGCAAGAGCA 259

Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGAAATGGCAAGAACAAATGTGCCAAGGTGAAATATCTCTCAAAAGAAATGTTG 319
Db 21 AsnTyrGlnGlnGlnGlyLysAsnAsnValProArgLeuIysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAATGTGATCATTTCATCGGTGGCCCAACAGCTCCAGTTATCATACCTTC 379
Db 41 GluSerAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyrHisThrPhe 60
QY 380 CTTTGGTGGAGGACGAGTGTAGCTGTATGCTTGGACCAAGGATCACAATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGGCCAGTATCTTACACCAAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATCAATGCAAGTGGGTGGAAAGACATCTCTGAAAGAAATGTGCTAAATTTTCATCAAGGTA 559
Db 101 AspGluCysIleTyrAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATATCAGACTCTGTACGCTGTGGAACGGGGCTTTTCATCCAAAT 619
Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTTGAAATTTGGACATCATCTGAGGACAAATATTTTAAAGCTCGAAGACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTGAAAAACGGCGTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGAGAAATATATCTCTGGAACCTGACACTGATTTTATGGGCGAGACTTTGCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCCGAATCTTGGGCACACCCCAATCAGGACAGACAGCATGATTTCCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCATATGATCAAAAGTTTCATTAGTCCCACCTCATCTCAGAGAGTGACAAATCTCGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATACTTTTCTCCGTGAAATGCAATAGATGAGAACACTCTGGAAAAGCT 979
Db 241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCACGCTAGATAGTTCAGATATGCAAGATGACTTTGGAGGGCCACAGAGTCTGGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGCAACATTCCTCAAGAGCTCGTCTGATTTGCTCAGTGCCAGGTCCAAATGGC 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGTATGAACTGACAGATGTATCTTAATGAATCTTAAAGATCTTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTCTGTATATGAGTGTTCACGACTTCCAGTAAACATTTTCAAGGATCAGCCGTG 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGAGTGTATGTGAGAGGGTGTCTCTTCCGTCCATATGCCCACAGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCCAACTATCAATGGGTGCTTATCAAGGAGAGTCCCTTATCCACGCCAGGAAC 1339

Db 361 GlyProAsnTyrGlnTyrValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCAGCAACAACTTTGGTGGTTTGGACTCTACAAAGGACCTTCCTGATGATCTTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspValIle 400
QY 1400 ACCTTTGCAGAAGTGCATCCAGCCATGTACAATCCAGTGTTCCTATGAACAATCCGCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTATCAAAACCGATTAATATCAATTTACAAATTTCTGTAGACCGAGTGGAT 1519
Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 CGAGAGATGACAGTATGATCTTATGTTATCGGAACAGATGTTGGGACCGTCTCTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAGAGGTTCTGTGGAAGAAATG 1639
Db 461 ValValSerIleProLysGluThrTyrAspLeuGluValLeuLeuGluMet 480
QY 1640 ACAGTTTTCGGAACCGACTCTATTTCCAGCAATGGAGCTTTCACTAAGCAGCAACAA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATGTTCAACGGCTGGGTTGCCAGCTCCCTTTACACGGGTGATATTATAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGAAGAAGCGTGTGAGTGTTCCTCGCCCGAGACCTTACTGTGCTGGGATGGTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCCACTGCAAGAGACGCCAAGACCAAGATATAAGAAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgGlnArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTGCTTACAGCTTACACCATGATTAATCACCATGGCCACAGCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisGlyHisSerPro 580
QY 1940 GAAGAGAGAAATCATCTATCGTGTAGAGATAGTACACATTTTTCGATGCGAGTCCGAG 1999
Db 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TCCAGAGACCGCTGCTATTGCGCAATTCAGAGCGCAATGAAGAGCAAAAGAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTGGATGATCATATCATCAGGACAGATCAGGCGCTTCTGCTAGTGTCTACAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGGATTCAGGCAATTTACCTCGCCATGCGGTGGAACATGGGTTCATACAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAAACCTGGAAGTCAATTCAGACAGCAATTTGGAAGAACTTCTTCATAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
QY 2240 GATGATGGAGTGGCTCTAAGACCAAGAAATGTCMAATAGCATGACACCTAGCCAGAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGTGTACAGAGACTTCATGACAGTTCATCAACCCCACTCTCAACACAGATGGATGAG 2359
Db 701 ValTyrTrpArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAAGTTTGGAAAGGGACCGAAACACAGCTCGGCAAGCGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740

QY 2420 CCAGGCAACAGTAACAATGGAAGCAGCTTACAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCCAGCAATTTGAGAGGGCACCCAGGAGTGTCT 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 3

US-08-488-935-2
; Sequence 2, Application US/08488935
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: Semaphorins
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RICHARD ARON OSMAN, Ph.D.
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,935
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/489,057
; FILING DATE: 09-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-2/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-488-935-2

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 8 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-488-935-2 (1-771)

QY 200 ATGGGCTGGTTAACTAGAGATTGTCTGCTTTTCTGGGAGATATTCTTACAGCAAGACA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGATCGGAGACAATGTCACAGGCTGAATTTATCTCAAAAGAAATGTTG 319
Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40

320 GAATCCAACTATGATCACTTTCAATGGCTTGGCCACAGCTCCAGTTATCATCTTC 379
Db |||||SerAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
380 CTTTTCGATGAGAAACGAGTAGCTGTATGTTGGAGCAAGGATCACATATTTTCATTC 439
Db |||||LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
440 GACCTGGTTAATATCAAGGATTTCAAAGATTTGTGTGCCAGTATCTTTACACCAAGA 499
Qy |||||AspLeuValAsnIleLysAspPheGlnLysIleValTyrProValSerTyrThrArg 100
500 GATGAATGCAAGTGGCTGGAAAGACATCTCGAAAGATGCTAAATTTTCATCAAGTA 559
Qy |||||AspGluCysLysTyrAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
560 CTTAAGGCATATAATCAGACTCACCTGTACGGCTGTGGAAACGGGGCTTTTCATCAATT 619
Qy |||||LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
620 TGCACCTACATTAATGGAATCATCTCTGAGCAATATTTTAAAGCTGGAGAACTCA 679
Db |||||CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
680 CATTTTGAACCGCGCTGGAGAGTCCATATGACCCCTAAGCTCTGACAGCATCCCTT 739
Qy |||||HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
740 TTAATAGATGAGAAATTAATCTGGAAGTCTGAGTGTATTTTATGGGGAGACTTTCCT 799
Qy |||||LeuIleAspGlyLeuLysSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
800 ATCTTCCGAACCTTTGGGCACACCCCAATCAGGACAGACAGCATGATTCAGAGTGG 859
Db |||||IlePheArgThrLeuGlyHisHisIleProIleArgThrGluGlnHisAspSerArgTyr 220
860 CTCATGATCCAAAGTTCAATAGTCCACCTCATCTCGAGAGTGACAACTCTCAAGAT 919
Qy |||||LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProLysVal 240
920 GACAAAGTATATCTTTCTTCGTCGAAATGCAATGATGAGAGACACTCTGGAAAGCT 979
Db |||||AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
980 ACTCAGCTAGTAATAGTGCAGATATGCAAGATGACTTTGGAGGCGCACAGAGTCTGGT 1039
Qy |||||ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
1040 AATAAATGCAACATTCCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCCTCAATGGC 1099
Qy |||||AsnLysTyrThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
1100 ATTGACACTCATTTTGTGAACTGCTGAGGATGATTCCTTAATCAACTTTAAAGATCTTAA 1159
Qy |||||IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
1160 AATCCAGTTGTATGAGTGTGTACGACTTCCAGTAACATTTTCAAGGATCAGCCGCTG 1219
Db |||||AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
1220 TGATGATAGCATGAGTGTGAGAGGGTGTTCCTTGTGTCATATGCCACAGGAT 1279
Qy |||||CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
1280 GGACCCAACTATCAATGGTGGCTTATCAAGGAGAGTCCCTATCCACGGCCAGGACT 1339
Qy |||||GlyProAsnTyrGlnTyrValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
1340 TGTCCAGCAAAACATTTGGTGGTTTTCACCTACAAAGGACCTTCCTGATGATGTTATA 1399
Qy |||||CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
1400 ACCTTTGCAGAAGTATCCAGCCCATGTACAATCCAGTGTTCCTATGAACAATCGCCCA 1459

401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
Qy |||||ATAGTGATCAAAACGAGTGTAAATTTATCAATTTTACAAAAATGCTGTAGACCGAGTGGAT 1519
Db |||||IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
1520 GCAGAAGATGCACAGTATGATGTTATGTTTATCGGAACAGATGTTGGGACCGTCTTTAAA 1579
Qy |||||AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
1580 GTAGTTTCAATTCCTAAGGACACTTGTATGATTTAGAGAGGTTCCTCTGGAAGAAATG 1639
Qy |||||ValValSerIleProLysGluThrTyrAspLeuGluGluValLeuLeuGluGluMet 480
1640 ACAGTTTTTTCGGGAACCGACTGCTATTTTCAGCAATGAGGCTTCCACTAAGCAGCAACA 1699
Qy |||||ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
1700 CTATATATTGTTCAACGGCTGGGTTGCCAGCTCCCTTTACACCGGTGTGATATTTAC 1759
Db |||||LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
1760 CGGAAGCGTGTCTGAGTGTGCTCGCCGAGACCTTACTGTGCTTGGATGGTCTTCT 1819
Qy |||||GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTyrAspGlySer 540
1820 GCATGTTCTCGCTATTTTCCACTGCAAGAGACGCAACAGCAGACAGATATAGAAT 1879
Qy |||||AlaCysSerArgTyrPheProThrAlaLysArgThrArgArgLysPheIleArgAsn 560
1880 GGAGACCCACTGACTCTGTTTCAGACTTACACCATGATATACCATGGCCACAGCCCT 1939
Db |||||GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisGlyHisSerPro 580
1940 GAAGAGAAATCATCTATGTTGTTAGAAATAGTAGACATTTTGGAAATGAGTCCGAG 1999
Qy |||||GluGluArgIleIleTyrGlyValGluAsnSerThrPheLeuGluCysSerProLys 600
2000 TCGCAGAGAGCGCTGTTCTATTGGCAATTCAGAGGCGAAATGAAGCGCAAAAGAAGAG 2059
Db |||||SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgGlyGlu 620
2060 ATCAGATGGATGATCATATCATCAGGACAGATCAAGCGCTTCTGCTACGTAGTCTACAA 2119
Db |||||IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
2120 CAGAAGATTCAGCAATTTACTCTGCTGATCGGTGGAAACATGGGTTCATACAACTCTT 2179
Qy |||||GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
2180 CTTAAGGTAAACCTGGAGTTCATTGACACAGCATTTGGAGAACTTCTTCATAAAGAT 2239
Qy |||||LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuHisLysAsp 680
2240 GATGATGAGATGCTCTTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAGAG 2299
Db |||||AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
2300 GTCTGGTACAGACTTCATGCTAGCTCATCAACACCCCAATCTCAACACATGGATGAG 2359
Qy |||||ValTyrTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
2360 TTCTGTCAACAAGTTTTCGAAAAAGGACCGAAAAACAACGTCCGCAAGCCAGGACATACC 2419
Qy |||||PheCysGluGlnValTyrLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
2420 CCAGGGAACAGTAACAAATGGAAGCACTTACAGAAAAATGAAGAGGTAGAAAAAGAGG 2479
Qy |||||ProGlyAsnSerAsnLysTyrLysHisLeuGlnLysLysLysGlyArgAsnArgArg 760
2480 ACCCAAGAAATTTGAGAGGCGACCCAGGAGTGTCT 2512
Qy |||||

```

Db      761 ThrHisGluPheGluArgAlaProArgSerVal 771
RESULT 4
US-08-489-057-2
; Sequence 2, Application US/08489057
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David R.
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: Semaphorins
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSES: RICHARD ARON OSMAN, Ph.D.
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/489,057
; FILING DATE: 09-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-2/RAO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-489-057-2
Alignment Scores:
Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 8 Gaps: 0
US-09-774-490-1 (1-2709) x US-08-489-057-2 (1-771)
QY      200 ATGGGCTGTTAACTAGGATTCCTGCTTTCTGGGAGTATTACTACAGCAAGACA 259
Db      1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY      260 AACTATCAGAAATGGGAAGAACAAATGTGCCAAGGCTGAAATATTCCTACAAAGAAATGTTG 319
Db      21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY      320 GAATCCAAACATGTGATCACTTTCAATGGCTTGGCCAAAGCTCAGATTATCATCTTC 379
Db      41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
QY      380 CTTTGGATGAGCAAGGAGTAGGCTGTATGTTGGAGCAAGGATCACATATTTTCATTC 439
Db      61 LeuLeuAspGluLysArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY      440 GACCTGGTTTAAATATCAAGGATTTTCAAAAGATTGTGTGCCAGTATCTTTACACCAAGA 499

```

```

Db      81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY      500 GATCAATCCAGTGGGCTGGAAAGACATCTGAAAGATGTGCTAAATTTTCATCAAGTA 559
Db      101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY      560 CTTAAGGCATATAATCAGACTCAGCTTGTACCCCTGTGGAACGGGGCTTTTCATCAAT 619
Db      121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY      620 TGCACCTACATTGAAATTGGACATCATCTCAGAGCAATATTTTAACTGGAGAACTCA 679
Db      141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY      680 CATTTTGAAACGGCGCTGGGAAGAGTCCATATGACCTTAGCTGCTCAGAGCATCCCTT 739
Db      161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY      740 TTAATAGATGGAGAAATTATCTCTGGAACTCAGCTGATTTTATGGGCGGAGACTTTGCT 799
Db      181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY      800 ATCTTCCGAACCTCTGGGACACCAACCAATCAGGACAGAGCATGATTCAGGTGG 859
Db      201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY      860 CTCATGATCCAAAGTTTCATTAGTCCACCTCATCTCAGAGAGTGACAATCCCTCAAGAT 919
Db      221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY      920 GACAAAGTATATCTTTCTCCGTGAAATCAATAGATGGAGAACACCTCTGGAAAAGCT 979
Db      241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY      980 ACTCACGCTAGAATAGGTGAGATATGCAAGAAATGACTTTGGAGGSCACAGAGTCTGTG 1039
Db      261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY      1040 AATTAATGGACAACATCTCCARAGCTGCTGATTTGCTCAGTCCAGGTCCTCAATGCG 1099
Db      281 AsnLysTrpThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY      1100 ATTGACACTCATTTTGTGATGAATCGAGGATGTATTCTTAATCAACTTTTAAAGATCCTAAA 1159
Db      301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY      1160 AATCCAGTTGATATGAGGTGTTTACGACTTCCAGTAAACATTTTCAAGGATCAGCCGTG 1219
Db      321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY      1220 TGTATGTATAGCATGAGTGTGAGAGGGTGTCTTGTGCTCATATGCCACACAGGAT 1279
Db      341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY      1280 GGACCCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCACGGCCAGGAAC 1339
Db      361 GlyProAsnTyrGlnTrpValProTyrGlnArgValProTyrProArgProGlyThr 380
QY      1340 TGTCCCGAGAAAACATTTGGTGGTTTGTACCTCACAAGGACCTTCTCGATGATGTTATA 1399
Db      381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY      1400 ACCTTTCCAGCAAGCTCATCCAGCCATGTACAAATCCAGTGTTCCTATCAACAATCGCCCA 1459
Db      401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetCAsnAsnArgPro 420
QY      1460 ATAGTGATCAAAACCGGATGTAATTTATCAATTTTACAAAATTTGCTGAGACCGAGTGGAT 1519
Db      421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY      1520 GCAGAGATGACAGTATGATGTTTATCGGAACAGATGTTGGACCGCTCTTAAA 1579
Db      441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460

```


QY 980 ACTCAGCTAGATAGTGCAGATATGCAAGATGACTTTGGAGGCGCAGAGTCTGCTG 1039
DB 261 ThrHisAlaAargileGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGACAACTTCCTCAAGCTCTCTGATTGCTCAGTGCCAGGTCCAAATGCG 1099
DB 281 AsnLysTrpThrPheLeuLysAlaAargLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATGCACTCATTTTGTAGAACTGCAGGATGTATTCCTTAATGAATTTAAAGATCCTAAA 1159
DB 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGTATATGAGTGTATACGACTCCAGTCAATTTTCAAGGGATCAGCCGTG 1219
DB 321 AsnProValValTyGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGATGATGATGAGAGGGTGTTCCTTGGTCCATATGCCCCACAGGAT 1279
DB 341 CysMetTySerMetSerAspValAargValPheLeuGlyProTyAlaHisArgAsp 360
QY 1280 GGACCACTATCAATGGTGCTTATCAAGGAAGTCCCTATCCACGGCCAGAACT 1339
DB 361 GlyProAsnTyGlnTrpValProTyGlnGlyA-gValProTyProArgProGlyThr 380
QY 1340 TGTCCTCCAGCAACATTTGGTGGTTTGNACTCTACAAAGGACCTTCTCGATGATGTTATA 1399
DB 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAAGAAGTCAATCCAGCCATGTACAAATCCAGTGTTCCTATGAACAAATCGCCA 1459
DB 401 ThrPheAlaAargSerHisProAlaMetTyAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACGGATTAATATCAATTTACAAATTTGCTGAGACCGAGTGTAT 1519
DB 421 IleValIleLysThrAspValAsnTyGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 CGAAGATGGACAGTATGATGTTATGTTATCGGAACAGATGTTGGACCGTCTTTAAA 1579
DB 441 AlaGluAspGlyGlnTyAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAAGAAATG 1639
DB 461 ValValSerIleProLysGluThrTrpTyAspLeuGluValLeuLeuGluMet 480
QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTCCAGAAATGAGACTTTCACATAGCAGCAACA 1699
DB 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATGTTCAACGCTGGGTGTCGCCAGCTCCCTTTACACCGGTGTGATATTAC 1759
DB 501 LeuTyIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTy 520
QY 1760 GCGAAGCGGTGCTGAGTGTTCCTCGCCCGAGACCTTACTGTGCTTGGAGTGTCT 1819
DB 521 GlyLysAlaCysAlaGluCysCysLeuAlaAargAspProTyCysAlaTrpAspGlySer 540
QY 1820 GATGTTCTCGCTATTTCCCACTGCAAGAGACGCAAGACGACAGATATAAGAAAT 1879
DB 541 AlaCysSerArgTyPheProThrAlaLysAargThrArgGlnAspIleAargAsn 560
QY 1880 GGAGACCCACTCACTGTTTCAGACTTACACCATGATATACCATGGCCACAGCCCT 1939
DB 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisIleGlyHisSerPro 580
QY 1940 GAAGAGAAATCATCTATGCTGTAGAGATAGTAGACATTTTGGAAATGCAAGTCCGAG 1999
DB 581 GluGluAargIleIleTyGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TGCAGAGAGCCCTGCTATTGGCAATTCAGAGCGGAAATGAAGAGCGGAAAGAGAG 2059
DB 601 SerGlnAargAlaLeuValTyTrpGlnPheGlnAargAsgAsnGluAargLysGluGlu 620

QY 2060 ATCAGATGGATGATCATATCATCAGACAGATCAAGCCCTTCTGCTAGTGTACAA 2119
DB 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGATTACCAATTACCTCTGCATCGGTGGAAACATGGGTTTCATACAACTCTT 2179
DB 641 GlnLysAspSerGlyAsnTyLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAACTCGGAGTCAATTCACACAGACATTTGGAGAACTTCTTCATAAAGAT 2239
DB 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuHisLysAsp 680
QY 2240 GATCATCGAGATGCTCTTAAGACCAAAAGAAATGTCCAAATAGCATCAGACACCTAGCCAGAG 2299
DB 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGTACAGAGACTTCATGCAGCTCATCAACCAACCCCAATCTCAACACATGGATGAG 2359
DB 701 ValTrpTyArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTCAACAAGTTTGGAAAAGGACCGAAAACAACGTCGCGCAAGCCAGGACATACC 2419
DB 721 PheCysGluGlnValTrpLysArgAspArgLysGlnAargGlnAargProGlyHisThr 740
QY 2420 CCAGGGAACAGTAACAATGAAGCACTTACAGAAATAAGAAAGGTAGAAAACAGGAGG 2479
DB 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnLysLysGlyArgAsnAargArg 760
QY 2480 ACCCAAGAAATTTAGAGGCGACCCAGGAGTGC 2512
DB 761 ThrHisGluPheGluAargAlaProArgSerVal 771

RESULT 6
US-09-791-537-31106
; Sequence 31106, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791.537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 31106
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-31106

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-791-537-31106 (1-771)

QY 200 ATGGGCTGTTAACTAGATTTCTGTCTTTCTGGGAGATTAATCTACAGCAAGACA 259
DB 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaAargAla 20
QY 260 AACTATCAGATCGGAAGAACAAATGTCCAGGCTGAAATTTATCTCAAAAGAAATGTTG 319
DB 21 AsnTyGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyLysGluMetLeu 40
QY 320 GAATCCAAATATGATCATTTCATCGGTGGCCCAACAGCTCCAGTTATACCTTC 379
DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyHisThrPhe 60

QY 380 CTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAGGATCACATATTTTCATTC 439
Db |||||
61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGGCCAGTATCTTACACCAAGA 499
Db |||||
81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATGAATGCAAGTGGCTGGAAAAGACATCTCTGAAAGAAATGTGCTAAATTTTCATCAAGTA 559
Db |||||
101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCACCTGTACGCTGTGGAACGGGGCTTTTCATCCAATT 619
Db |||||
121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATGGAATTGGACATCATCTGAGGCAATATTTTAAAGTGGAGAACTCA 679
Db |||||
141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAAAACGGCGTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTT 739
Db |||||
161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGGAATTTACTCTCGAACTGCAGCTGATTTTATGGGCGAGACTTTCCT 799
Db |||||
181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCCGAACCTTGGGACACACCACCCATCAGGACAGGACGATGATCCAGGTGG 859
Db |||||
201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCATATGATCAAAAGTTCATTAGTCCCACTCATCTCAGAGAGTGAATCCTCAAGAT 919
Db |||||
221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTTCTCGTGAAATGCAATAGATGGAGAACACTCTGGAAAAGCT 979
Db |||||
241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGATAGGTGAGTATGCAAGATGATCTTGGAGGCCACAGAGTCTGTG 1039
Db |||||
261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGAACAATCTCTCAAGCTCGTCTGATTGCTCAGTCCAGGTCCCAATGSC 1099
Db |||||
281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCAATTTTGTGAATCTGAGGATGTATTCCTAATGAATTTTAAAGATCTTAAA 1159
Db |||||
301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTGTATATGGAGTGTTCAGACTCCAGTATCAATTTTCAAGGATCAGCCGTG 1219
Db |||||
321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGAGTGTGAGAAGGTGTTCTTGTGTCATATGCCCCACAGAGGAT 1279
Db |||||
341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCAACTATCAATGGGTGCCTTATCAAGGAAGAGTCCCTATFCCACGGCCAGGAAT 1339
Db |||||
361 GlyProAsnTyrGlnTrpValProTyrGlnArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCCAACAAACATTTGGTGGTTTTCATCTACAAAGAGACCTTCTCTGTATGATGTATA 1399
Db |||||
381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAAGAAGTCAATCCAGCATGATCAATCCAGTGTTCCTATGAAACAATGCCCA 1459
Db |||||
401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420

RESULT 7

QY 1460 ATAGTGATCAAAACGGATGTAAATATCAATTTACAAATTTGCTGTAGACCGAGTGGAT 1519
Db |||||
421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAGATGCACAGTATGATGTTATCGGNACAGATGTGGGACCGTCTCTAAA 1579
Db |||||
441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAGAGGTTCTCTCGAAGAAATG 1639
Db |||||
461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTTCGGAAACCGACTCTATTTTACAGCAATGGAGCTTTCACATTAAGCAGCAAA 1699
Db |||||
481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
QY 1700 CTATATATTGTTTCAACGGCTGGGTTGCCAGCTCCCTTTACACCGGTGTGATATTTAC 1759
Db |||||
501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GCGAAAGCTGCTGCTGAGTGTGCTGCCCGAGACCCCTTACTGTGCTTGGATGGTTCT 1819
Db |||||
521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCACTGCAAGAGACGCAACAGCAAGATATAGAAT 1879
Db |||||
541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgAspIleArgAsn 560
QY 1880 GGAGACCACTGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1939
Db |||||
561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580
QY 1940 GAAAGAGAAATCATCTATGTTAGAGAAATAGTAGACATTTTGGAAATGAGTCCGAG 1999
Db |||||
581 GluGluArgIleIleTyrGlyValGluAsnSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGAGCGCTGCTGCTATTTGGCAATTCAGAGGCGAAATGAAGCGCAAGAGAGAG 2059
Db |||||
601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAGTCTACAA 2119
Db |||||
621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGGATTCAGGCAATTTACTCTGCTCATCGGTGGAACATGGTTCATACAAACTCTT 2179
Db |||||
641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAAACCTTGGAGTCAATTGACACAGAGCATTTGGAGAACTTCTTTCATAAAGAT 2239
Db |||||
661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
QY 2240 GATCATGGAGTGGCTCTAAGACCAAGAAATGTCATAGCATGACACCTTAGCCAGAG 2299
Db |||||
681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGGTACAGAGACTTTCAGTCACTCATCAACCAACCCCAATCTCAACACCATGATGATCAG 2359
Db |||||
701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAAGTTTCGAAAAGGACCGGAAACCAACCGTCGGCAAGCCAGGACATACC 2419
Db |||||
721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGAACAGTAACTAATGAGACCTTACAGAAATTAAGAAAGTGAAGAACAGGAGG 2479
Db |||||
741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCCACCAGTTTGAAGGGCACCCAGGAGTGTCT 2512
Db |||||
761 ThrHisGluPheGluArgAlaProArgSerVal 771

US-10-067-632-54
; Sequence 54, Application US/10067632
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; Kolodkin, Alex L.
; Matthes, David
; Bentley, David R.
; O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/067,632
; FILING DATE: 04-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,610
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/835,268
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-10-067-632-54

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 26 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-067-632-54 (1-771)

QY 200 ATGGCTGTTAACTAGGATTCCTGCTTTCTGGGAGTATTAATCTACAGCAAGACA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGAAATGGGAGCAATGTGCCAAGCTGAAATTAATCTACAAAGAAATGTG 319
Db 21 AsnTyrGlnAsnGlyLysAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAATGTGATCACTTTCAATGGCTTGGCCAAAGCTCCAGTTATCATACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
QY 380 CTTTGGATGAGCAACGGAGTAGGCTGTATGTGGAGCAAGGATCAATATTTTCATTC 439
Db 61 LeuLeuAspGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80

QY 440 GACCTGTTAATATCAAGGATTTTCAAAGATTTGTGTGCCAGTATCTTACACCAAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArg 100
QY 500 GATGAATCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTCTAAATTTTCATCAAGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCATTGTACCCCTGTGGAAACGGGGCTTTTCATCAATT 619
Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTGAAATTTGGACATCATCTCAGGACAAATATTTTAACTCGAGACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisProGlnAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAACACGGCTGGGAAGAGTCCATATACCCCTAAGCTCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGAGAAATTAATCTCTGGAACCTCAGCTGATTTTATGGGCGAGACTTTC 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCCGAACTCTTGGSCACCCACCATCAAGGACAGACAGCAGCATGATTCCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCAATGATCCAAAGTTTCATTAGTCCCACTCATCTCAGAGAGTGACAAATCTCTGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATCTTTTCTCCGTGAAATTCATATAGATGAGAGACATCTCTGGAAGCT 979
Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGATAGGTGAGATATCAAGAAATCACTTTTGGAGGCGACAGAAAGTCTGGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGCAACATTCCTCAAGCTCGTCTGATTTTCTCAGTCCCAAGGTCCCAATGCG 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGTAGTAACTCAGAGTGTATTCCTAATGAATTTTAAAGATCTTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGTATATGGAGTGTTCACGACTTCCAGTAACATTTTCAAGGGATCAGCCGTG 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGAGTGTGAGAAAGGTGTCTTGTGTCCATATGCCACAGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCAACTATCAATGGTGCCTTATCAAGGAGAGTCCCTATCCACGCCCAGGAAT 1339
Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCCAAGAAAACATTTGGTGGTTTGTACTCTCAAGAGGACCTTCTCTGATGATGTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAsnSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTCCAGAAAGTCAATCCAGCCATGTACAAATCCAGTGTTCCTATGAAACAATGCCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGTCAAAACGGATGTAATTAATCAATTTACAAAATTCGTAGACCGAGTGGAT 1519
Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440


```
QY 1520 GCAGAGATGCGACGATATGATGTTATGTTATTCGGAACAGATGTTGGGACCGTCTTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGCAGACTGCTATGATTTTGAAGAGAGTCTGCTGGAAGAAATG 1639
Db 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluMet 480
QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTCCAGCAATGGAGCTTTCCACTAAGCAGCAACAA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATTGTTCAACGCTGGGTTGCCAGCTCCCTTTACCGGTGTATATTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuLeuHisArgCysAspIleTyr 520
QY 1760 GCGAAGACGCTGCTGAGTGTTCCTCGCCGAGACCCCTTACTGTGCTGGGATGTTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCCACTGCAAGAGACGCGACAGCAAGATATAAGAAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgThrArgArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTCAGCTGTTCCAGCTTACACCATGATTAATCACCAGCCGCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisGlyHisSerPro 580
QY 1940 GAAGAGAGAAATCATCTATGCTGTAGAGATAGTACACATTTTGGATGCGAGTCCGAG 1999
Db 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGACGCTGCTGTTATTTGGCAATTCAGAGCGCAAAATGCAAGAGCGAAAGAAG 2059
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTGGATGATCATATCATCAGACAGATCAAGGCTTCTGCTAGTGTACAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGGATTTCAGCAATTAACCTCTGCCATGGGTGGAACATGGGTTTCATACAAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAACCTCGAAGTATTGACACAGAGCATTTTGGAAAGAACTTCTTCATAAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
QY 2240 GATGATGGAGATGGCTCTAAGACCAAGAAATGTCAATAGCATGACACCTAGCCAGAAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGGTACAGAGCTTCATGACGCTCATCAACCCACCCCAATCTCAACAGCATGATGAG 2359
Db 701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAGTTTGGAAAGGACCGAAACCAACGTCGCAAGGCGCAGGACATACC 2419
Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGACAGTAACAATGGAGCACTTCAAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCCAGCAATTTGAGAGGCGCCACCGAGGTGTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 8
US-10-097-340-284
; Sequence 284, Application US/10097340
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
```

```
; APPLICANT: Manjula GANNAVAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 284
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-284

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 26 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-097-340-284 (1-771)

QY 200 ATGGGCTGGTTAACTAGGATTGTCTGCTTTCTGGGAGTATTACTTACAGCAAGACA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGAATGGAAGAACAAATGTGCCAAGGCTGAAATTTATCTCAAAAGAAATGTTG 319
Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAACAAATGTGATCAGCTTTCAATGGCTTGGCCAGCTCCAGTTCATACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyrHisThrPhe 60
QY 380 CTTTGTGATGAGGAACCGAGTAGCTGTATGTTGGAGCAAGGATCAGATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACACCAAGAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrTrpArgArg 100
QY 500 GATGAATGCAAGTGGGCTGGGAAAGACATCCTGAAAGAATGTGCTAATTTTCATCAAGGTA 559
```

Db 101 AspGluCysLeuThrAlaGlyLysAspLeuLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCATCTGTACGGCTGTGGAAACGGGGCTTTTCATCCAATT 619
Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTAATGGACATCATCTGTAGGACATATTTTAAAGCTGGAGAACTCA 679
Db 141 CysThrTyrIleGluLeuGlyHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAAGACGGCGTGGGAAGAGTCCATATGACCCCTAAGCTGCTCAGACATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGGAATATTAATCTGTGAAGTGTGAGTGTATTTATGGGGGAGACTTTGCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCGAACCTTTGGGCACCAACCAATCAGACAGACAGACATGATTCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTyr 220
QY 860 CTCAATGATCCAAAGTTCATTAGTGCACCTCATCTCAGAGAGTGACAAATCCTGAAGT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTCTTCGCGTGAATGCAATGCAATAGATGGAGACACTCTGGAAGCT 979
Db 241 AspyLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGATAGTGCAGATGTCAGATGACTTTGGAGGACACAGAGTCTGGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyHisArgSerLeuVal 280
QY 1040 AATAATGACAACTTCCTCAAGCTCGTCTGATTGCTCAGTGCCAGGTCCAAATGGC 1099
Db 281 AsnLysTyrThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTGATGAACATGACGAGATGATTCCTTAATGAATCTTAAAGATCCTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCAGTTGTATATGGAGTGTTCAGCTTCAGTAACTTTCAAGGATCACCGCTG 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGATGTGAGAGGGTTCCTTTGGTTCATATGCCACACAGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCAACTATCAATGGGTGCTTATCAAGGAAGTCCCTATCCAGGGCAGAACT 1339
Db 361 GlyProAsnTyrGlnTyrValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCAGCAAACTTTGGTGGTGTGACTCTACAAAGGACCTTCCTGATGATCTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAAGAAAGTCCAGCCCATGTACAAATCCAGTGTTCCTATGAACAAATCCGCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTATCAAAACGGATTAATATCAATTTACAAATTTGCTGTAGACCGAGTGGAT 1519
Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 CGAGAAGTGGACATGATGATCTTATGTTTATCGGAACAGATGTTGGGACCGTCTTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAAGAAATG 1639

Db 461 ValValSerIleProLysGluThrTyrAspLeuGluGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTCACCAATGGAGCTTCACATAAGCAGCAACA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATTTGGTTCAAACGGCTGGGTGCCAGCTCCCTTTTACACCGGTGTGATATTTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GCGAAACGGTGTCTGAGTGTTCCTCCCGAGACCTTACTGTGCTTGGATGGTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTyrAspGlySer 540
QY 1820 GCATGTTCTCGTATTTCCTTCCACTCCAAAGAGACGCAACAGCAGACAGATATAGAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgThrArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTCTGTTTCAGACTTCACCATGATATACCATGGCCACAGCCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580
QY 1940 GAAGAGAGATTCATCTATGTTGTAGAGATAGTACACATTTTGGAAATCAGTCCGAG 1999
Db 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGAGCGCTGCTTATTCGCAATTCAGAGGCGAATGAAGCGGAAAGAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTyrGlnPheGlnArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAGTCTACA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGGATTCAGCAATTAACCTCTGCATCGGTGGAACATGGGTTTCATACAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGTTAACCTTGGAGTCAATTCACACAGAGCATTTGGAGAACTTCTTTCATAAAGT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
QY 2240 GATCATGAGATGCTCTTAAGACCAAAAGAAATGCTCAATAGCATGACACCTAGCCAGAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGTACAGACTTCATGACGCTCATCAACCAACCCCAATCTCAACACGATGGATGAG 2359
Db 701 ValTyrTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGACCAAGTTTGGAAAGGCGGAAACCAACGTCGGCAAGGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTyrLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGCAACAGTAACTTGAAGCACTTCAAGAAATAAGAAAGGTAGAAACAGAGG 2479
Db 741 ProGlyAsnSerAsnLysTyrLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCCAGCAATTTAGAGGCGACCCAGGAGTCTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 9

US-10-247-671-164
; Sequence 164, Application US/10247671
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247.671

; CURRENT FILING DATE: 2002-09-18
 ; PRIORITY FILING DATE: 60/323,784
 ; PRIORITY FILING DATE: 2001-09-19
 ; NUMBER OF SEQ ID NOS: 186
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 164
 ; LENGTH: 771
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No: 1930967CD1
 US-10-247-671-164

Alignment Scores:
 Pred. No.: 0 Length: 771
 Score: 4201.00 Matches: 771
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 85.80% Indels: 0
 DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-247-671-164 (1-771)

QY 200 ATGGCTGTTAACTAGGATGCTGCTCTTCTGGGAGTATTACTTACAGCAAGACA 259
 Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaAlaArgAla 20
 QY 260 AACTATCAGATGGGAAGAACAACTGCGCAAGGCTGAATATCTCAACAAAGAAATGTTG 319
 Db 21 AsnTyrGlnAsnGlyLeuAsnValProArgLeuLeuLeuSerTyrGlyGluMetLeu 40
 QY 320 GAATCCAACTATGATCACTTCAATGGCTTGGCCAAAGCTCCAGTTCATCATCTTC 379
 Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
 QY 380 CTTTGGATGAGGAACGAGTAGGCTGATGTTGGAGCAAGAGTCAATATTTTCATTC 439
 Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisLeuPheSerPhe 80
 QY 440 GACCTGGTTAATATCAAGGATTTTCAAGATTTGTCGCCAGTATCTTACACCAAGAGA 499
 Db 81 AspLeuValAsnIleLeuAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
 QY 500 GATGAATCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGCTGTAATTTTCATCAAGGTA 559
 Db 101 AspGluCysLysTrpAlaGlyLeuAspLeuLysGluCysAlaAsnPheIleLysVal 120
 QY 560 CTTAAGGCATATATCAGACTCACTTGATCGCTGTGGAAAGCGGGCTTTTCATCCAAAT 619
 Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
 QY 620 TGCACCTACATTGAAATTTGGACATCATCTCAGGACAAATATTTTAACTCGGAGACTCA 679
 Db 141 CysThrTyrIleGluIleLeuIleHisProGluAspAsnIlePheLysLeuGluAsnSer 160
 QY 680 CATTTTGAAACGGCCGTGGGAAGAGTCCATATGACCCCTAAAGCTGCTCAGACGATCCCTT 739
 Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
 QY 740 TTAATAGATGAGAAATTAATCTCTGGAACTGAGCTGATTTTATGGGCGGAGACTTGCT 799
 Db 181 LeuLeuAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
 QY 800 ATCTTCCCACTCTGGGCACACACCCATCAGGACAGAGCAGCATGATTCAGGTGG 859
 Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
 QY 860 CTCAATGATCAAAAGTTTCATTAGTCCCACTCATCTCAGAGAGTGACAAATCTCGAAGAT 919
 Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
 QY 920 GACAAAGTATATCTTTCTTCCTCGTGAATAATGCAATAGATGGAGAACACTCTGGAAAGACT 979

Db 241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
 QY 980 ACTCAGCTAGAAATAGGTGAGATATGCAAGATGCTTTGGAGGCGCAGAGAGTCTGGTG 1039
 Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
 QY 1040 AATAAATGGACAACATCTCTCAAGCTCGTCTGATTGCTCAGTGCCAGGTCAAAATGCG 1099
 Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
 QY 1100 ATTGACACTCATTTTGTAGAACTGAGATGATATCTTAATGAACATTTAAAGATCCTAAA 1159
 Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
 QY 1160 AATCCAGTTGTATATGGAGTGTTCACGACTTCCAGTAACATTTTCAAGGATCAGCCGTG 1219
 Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
 QY 1220 TGTATGTATAGCATGAGTGTGAGAGGGTGTTCCTTGGTCCATATGCCACAGGGAT 1279
 Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
 QY 1280 GGACCAACTATCAATGGGTGCTTATCAAGAGAGTCCCTATCCAGGCGCAGAACT 1339
 Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
 QY 1340 TGTCCACAAAAACATTTGGTGGTTTGCACCTACAAAGGACCTTCTGTATGATGTTATA 1399
 Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
 QY 1400 ACCTTTTGAAGAAGTCAATCCAGCATGATCAATCCAGTGTTCCTTATCAACAAATGCCCA 1459
 Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
 QY 1460 ATAGTGATCAAAACGGATGTAATTTCAATTTTACAAATTTGCTGAGACCGAGTGGAT 1519
 Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
 QY 1520 GCAGAAGATGCACAGTATGATGTTATCGGACAGATGTTGGGACCGTCTCTTAAA 1579
 Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
 QY 1580 CTAGTTTCAATTCCTAAGGAGACTTTGGTATGATTTAGAGAGGTTCTCTCGAAGAAATG 1639
 Db 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluGluMet 480
 QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTTCACCTAAGCAGCAACA 1699
 Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
 QY 1700 CTATATATTGTTTCAACGGCTGGGTTGCCAGCTCCCTTTACACCGGTGTGATATTAC 1759
 Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
 QY 1760 GGGAAAGCGTGTGAGTGTGCTCGCCGAGACCTTACTGTCTGTGGATGGTCT 1819
 Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
 QY 1820 GCATGTTCTCGCTATTTTCCACTGCAAGAGAGCGCAAGACGCAAGATATAGAAT 1879
 Db 541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgGlnAspIleArgAsn 560
 QY 1880 GGAGACCCACTGACTCTGTTCCAGCTTACCATGATATCACTATGATGCGCCAGCCCT 1939
 Db 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580
 QY 1940 GAAAGAGAAATCATCTATGTTGTAAGAATAGTAGACATTTTGGAAATGAGTCCGAG 1999
 Db 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
 QY 2000 TCGCAGAGAGCGTGTGCTTATTTGGCAATTCAGAGGCGAAATGAAGAGCGGAAAAAGAG 2059

Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCCTTCTCTAGGTAGTCTACAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAGGATTCAGGCAATTACCTCTCGCCATCGGTGGAACTGGGTTCATACAAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuGlyHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAAACCTCGAAGTCAATGACACAGAGCAATTTGGAGAACTTCTTCATAAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuHisLysAsp 680
QY 2240 GATGATGAGATGGCTTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAGAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGGTACAGAGCTTCATGCGCTCATCAACACCCCAATCTCAACACGATGGATGAG 2359
Db 701 ValTyrTrpArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGCAAGATTTGGAAAAGGACCGAAACCAACGTCGCGCAAGGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTyrLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CAGGGACAGTAACAATAGGAAGCATTACAGAAATATAGAAAGGTAGAAACAGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCCAGCAATTTGAGAGGCGCCACCGAGGTGTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 10

US-10-262-538-10
; Sequence 10, Application US/10262538
; GENERAL INFORMATION:
; APPLICANT: Alitalo et al
; TITLE OF INVENTION: NEUROFILIN/VEGF-C/VEGFR-3 MATERIALS AND METHODS
; FILE REFERENCE: 28967/37564
; CURRENT APPLICATION NUMBER: US/10/262,538
; CURRENT FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-538-10

Alignment Scores:

Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-262-538-10 (1-771)

QY 200 ATGGCTGGTTAACTAGGATTCCTGCTTTCTGGGAGTATTACTTACAGCAAGACA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGAAATGGGAAGAACAAATGTGCCAAGCGTGAATATTCCTCAAGAAATGTGTG 319
Db 21 AsnTyrGlnAsnGlyLysAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAACAATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCATCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrThrPhe 60

QY 380 CTTTTCGATGAGGAACGAGTAGGTCTGTATCTTGGAGCAAGAGTACATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAATATCAAGGATTTTCAAAGANTGTGTGGCCAGTATCTTACACCAAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATCAATGCAAGTGGCTGGAAAGACATCTGGAAGNATGTCTAATTTTCATCAAGTGA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCACTGTACCCCTGTGGAACGGGGCTTTTCATCCAATT 619
Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTTGAAATTTGGACATCATCTCTGAGGACAAATATTTTAACTCGAGAACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAACCGCCCTGGGAAGTCCATATGACCTAAGCTGCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGAGAAATATATCTCTGGAACCTGACGTGATTTTATGGGCGGAGACTTTGCT 799
Db 181 LeuLeuAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCCGAACTCTTGGGACCCACCCCAATCAGGACAGAGCATGATTCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCAATGATCAAAAGTTCAATAGTCCACCTCATCTCAGAGAGTGCAATCCTGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTCTCCGTGAAATCAATAGATGAGAGAACACTCTCGGAAAAGCT 979
Db 241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGATAGTTCAGATATGCAAGAAATGACCTTTGGAGGCGACAGAGTCTGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGCAACATTCCTCAAGCTCGTCTGATTTGCTCAGTGCCAGGTCCCAATGSC 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGTGATGAACCTGAGGATGTATTTCTTAATGAACCTTTAAAGATCCTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGATATGAGTGTTCAGACTTCCAGTAAACATTTTCAAGGGATCAGCCGTG 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGATACATGATGATGTGAGAGGTGTCTCTTGGTCCATATGCCACAGGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCCACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTATCCACGGCCAGGAACT 1339
Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCCAAGAAAAATTGGTGGTTTGTACCTCTCAAGAGGACCTTCTCTGATGATGTTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTCGCAAGACTCATCCAGCCATGTACAAATCCAGTGTTCCTTATCAACAAATCGCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACCGATGTAATTTTACACAAATTTGTCGTAGACCGAGTGGAT 1519

Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTyr 220
QY 860 CTCATGATCCAAAGTTCTATTAGTCCACCTCATCTCAGAGAGTGAACAATCCCTGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTCTTCCTCGTCAAAATGCAATAGATGGAGAACACTCTCGGAAAAGCT 979
Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGATAGGTGAGATATGCAAGATGCTTTGGAGGCCACAGAAAGTCTGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGACACATCTCTCAAGCTCGTCTGATTCTCAGTCCAGGTCCAAATGCC 1099
Db 281 AsnLysTyrThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGATGAACCTGAGGATGTATTCTTAATGAATTTTAAAGATCCTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCAGTTGTATATGAGTGTATACGACTTCCAGTAAATTTCAAGGATCAGCCGTG 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGATGTGAGAGGTGTTCTTGTGTCATATGCCACAGGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCAACTATCAATGGTGCTTATCAAGGAAGAGTCCCTATCCACGGCCAGGAAT 1339
Db 361 GlyProAsnTyrGlnTyrValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCCAACAAACATTTGGTGGTTTGACTCTACAAAGGACCTTCTCGATGATGTTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAAGAGTCATCCAGCCATGTACATCCAGTGTTCCTATCAACAAATCGCCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACCGATGTAATTTATCAAAATTTACAAATTTGCTAGACCGAGTGGAT 1519
Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 CGAAGATGACAGATGATGTTATGTTATCGGAACAGATGTTGGACCGTCTCTTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGGAGCTTGGTATGATTTAGAGAGTTCTCTCGAAGAAATG 1639
Db 461 ValValSerIleProLysGluThrTyrAspLeuGluGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTTCGGAAACCGACTCTATTTACAGCAATGGAGCTTCCACTAAGCAGCAACA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CPTATATTTGGTTCAACCGCTGGGTGCTCCAGTCCCTTTACACCGGTGATATTTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTGGAGTGGTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTyrAspGlySer 540
QY 1820 GCATGTTCTCGTATTTTCCCACTGCAAGAGACCCACAAAGACCAAGATATAGAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn 560
QY 1880 GGAGACCACTGACTGCTGTTTCAGACTTACACCATGATATATCATCATGGCCACAGCCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580

QY 1940 GAAGAGAGATCATCTATGTGTAGCAATAGTAGCACATTTTGGATGCAGTCCGAAG 1999
Db 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGAGCGTGTCTATTGGCAATTCAGAGGCGAAATGAAGACGAAAAAGAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTyrGlnPheGlnArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTGATGATCATATCATCAGACAGATCAAGGCCTTCTGCTACGTAGTCTACA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGGATTGAGCAATTTACTCTGCGCATGCGGTGAAACATGGGTTCATACAAACTCT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAAACCTGGAGATCATGACAGAGCATTTGGAGAACTTTCTTCATAAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
QY 2240 GATCATGAGATGCTCTTAAGACCAAAAGAAATGTCCAAATAGCATGACACCTAGCCAGAAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGGTACAGAGACTTTCATGCAGCTCATCAACCCACCCCAATCTCAACACCATGGATGAG 2359
Db 701 ValTyrTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTCAACAAGTTTCGAAAAAGGACCGAAAAACAACGTCGGCAAGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTyrLysArgAspArgLysGlnArgArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGAACAGTAACAATCGAAGCACTTACAGAAATTAAGAAAGGTAGAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTyrLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCCAACCAATTTAGAGGCGACCCAGGAGTGTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771
RESULT 12
US-09-791-537-101922
; Sequence 101922. Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomax, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMEB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent In version 3.0
; SEQ ID NO 101922
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-101922
Alignment Scores:
Pred. No.: 0 Length: 772
Score: 4057.50 Matches: 740
Percent Similarity: 98.58% Conservative: 21
Best Local Similarity: 95.85% Mismatches: 10
Query Match: 82.87% Indels: 1
DB: Gaps: 1
US-09-774-490-1 (1-2709) x US-09-791-537-101922 (1-772)
QY 200 ATGGCGTGGTTAACTAGAGATTGTCTCTTTCTGGGGAGTATTACTTACAGCAGACCA 259
|||||

1 MetGlyTrpPheThrGlyIleAlaCysLeuPheTrpGlyIleLeuLeuThrAlaAlaGala 20
260 AACTATCAGAATGGGAAGAACAAATGTGCCAAGCGTGAATATATCTCAAGAAATGTTG 319
21 AsnTyrAlaAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
320 GAATCCAAATGTCATCACATTTCAATGGCTGGCCAAACAGCTCCAGTTATCATACCTTC 379
41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
380 CTTTGGATGAGGAACGAGTAGGTGTATGTGGAGCAAGAGGATCACATATTTTCATTC 439
61 LeuLeuAspGluLysSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
440 GACCTGGTTAATATCAAGCATTTTCAAAAGATTTGTGGCCAGTATCTTACACAGAAGA 499
81 AsnLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
500 GATGAATCAAGTGGCTGGAAAGAACACATCTGGAAGAAAGATGTCTAAATTCATCAGGTA 559
101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
560 CTTAAGGCATATATCAAGACTCACCTGTACGCTGTGGAACGGGGCTTTTCATCCAAAT 619
121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
620 TGCACCTACATGAAATTTGGACATCATCTCGAGGACAAATATTTTAACTGGAGAATCA 679
141 CysThrTyrIleGluValGlyHisIleProGluAspAsnIlePheLysLeuGlnAspSer 160
680 CATTTGAAACGGCGTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTT 739
161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
740 TTAATAGATGAGAAATTAATCTCGAACTCGAGCTGATTTATGGGCGAGACTTTGCT 799
181 LeuLeuAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
800 ATCTTCCGAATCTTTGGCCACCACCCATCAGGACAGAGCATGATTCAGAGTGG 859
201 IlePheArgThrLeuGlyHisIleHisIleProIleArgThrGluGlnHisAspSerArgTrp 220
860 CTCAATGATCAAAGTTCAATAGTCCACCTCATCTCAGAGAGTGACAATCTCGAAGAT 919
221 LeuAsnAspProArgPheIleSerAlaHisLeuIleProGluSerAspAsnProGluAsp 240
920 GACAAAGTATATCTTTCTCCGTGAAATGCAATAGATGAGACACACTCTGGAAAGCT 979
241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
980 ACTCAGCTAGATAGGTGAGATATGCAAGATGACATTTGGAGGCGCACAGAAGTCTGTG 1039
261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
1040 AATAAATGACAACTTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCCAAATGGC 1099
281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
1100 ATTGACACTATTTGATGAATCTCAGGATCTATTCTTAATGAATTTAAAGATCTTAAA 1159
301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnSerLysAspProLys 320
1160 AATCAGTGTATATGAGTGTTCACGACTTCCAGTAACTTTCAAGGGATCAGCGGTG 1219
321 AsnProIleValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
1220 TGTATGATACATGATGTGAGAAGGTGTTCTTGTGTCATATGCCACAGGAT 1279
341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
1280 GGACCCAACTATCAATGGTCCCTTATCAAGAGAGTCCCTTATCCACGCCAGGAACT 1339
361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr

1340 TGTCACGACAAACATTTGGTGGTTTGACTCTCAAGAGCACTTCTCTGATGATGTTATA 1399
381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
1400 ACCTTTGCAGAAGTCATCCAGCCATGTACAATCCAGTGTTCCTATGAACAATGCCCA 1459
401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProIleAsnAsnArgPro 420
1460 ATAGTGATCAAAACGGATGTAATATCAATTTACAAAATTTGCTGAGACCGAGTGAT 1519
421 IleMetIleLysThrAspValAsnTyrGlnPheThrGlnIleValIleAspArgValAsp 440
1520 GCAGAAGATGCACAGTATGATGTTATGTTATCCGACACAGATGTTGGACCGTCTTTAAA 1579
441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAGAGGTCTCTCTGAGAGAAATG 1639
461 ValValSerValProLysGluThrTrpHisAspLeuGluValLeuLeuGluGluMet 480
1640 ACAGTTTTCGGGAACCGACTGCTATTTTCAGCAATGGAGCTTTCACCTAACGACGACAAA 1699
481 ThrValPheArgGluProThrThrIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
1700 CTATATATTTGGTTCAACGGCTGGGGTTGCCAGCTCCCTTTACACCGGTGTGATATTTAC 1759
501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
1760 GGGAAACGGTGTGCTGAGTGTTCCTCGCCGAGACCCCTTACTGTGCTGGGATGGTTCT 1819
521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
1820 GCATGTTCTCGCTATTTCCACCTGCAAGAGAGCGACAGACGACAGATATAGAAT 1879
541 SerCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn 560
1880 GGAGACCACTGACTCTACTGTTTCAGACTTA---CACCATGATAATCACTACGTCACAGC 1936
561 GlyAspProLeuThrHisCysSerAspLeuGlnHisIleAspAsnHisIleGlyHisSer 580
1937 CCTCAAGAGAGATCATCTATGTTGTTAGAGATAGTAGACATTTTGGAAATGAGTCCG 1996
581 LeuGluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerPro 600
1997 AAGTCGACAGAGCGCTGCTATTGGCAATTCAGAGGCGAATGAAGACGCAAGAAAGAA 2056
601 LysSerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluAspArgLysGlu 620
2057 GAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGCGCTTCTGCTACGTAGTCTA 2116
621 GluIleArgValGlyAspHisIleIleArgThrGluGlnGlyLeuLeuLeuArgSerLeu 640
2117 CAACAGAAGATTCAGGCAATTTACCTCTGCCATCGGTGGAAACATGGGTTTCATCAAACT 2176
641 GlnLysLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheMetGlnThr 660
2177 CTTCTTAAGGTAAACCTGGGAAGTCATTGACACAGAGCATTTGGGAAGCACTTCTTCATAAA 2236
661 LeuLeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLys 680
2237 GATGATCATGAGATGGCTCTTAAGACCAAGAAATGTCCAAATAGCATGACACCTAGCCAG 2296
681 AspAspAspGlyAspGlySerLysThrLysGluMetSerSerSerMetThrProSerGln 700
2297 AAGTCTGTGACAGACTTTCATGACGCTCATCAACACCCCAATCTCAACACGATGGAT 2356
701 LysValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAsp 720
2357 GAGTCTGTCAACAGTTTCGAAAAGGACCGGAAACAAACGTCGGCAAGCCAGGACAT 2416
721 GluPheCysGluGlnValTrpLysArgAspArgLysGlnArgArgGlnArgProGlyHis 740


```
QY 2417 ACCCCAGGAAACAGTAACAATGGAGCAGCTTACAGAAATAGAAAGGTAGAAACAGG 2476
   ::: |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 741 SerGlnGlySerSerAsnLysTrpLysHisMetGlnGluSerLysLysGlyArgAsnArg 760

QY 2477 AGGACCCAGCAATTTGAGAGGCGACCCAGGAGTGTC 2512
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 761 ArgThrHisGluPheGluArgAlaProArgSerVal 772

RESULT 13
US-09-791-537-72548
; Sequence 72548, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 72548
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-72548

Alignment Scores:
Pred. No.: 0 Length: 772
Score: 4035.50 Matches: 737
Percent Similarity: 98.32% Conservative: 22
Best Local Similarity: 95.47% Mismatches: 12
Query Match: 82.42% Indels: 1
DB: 22 Gaps: 1

US-09-774-490-1 (1-2709) x US-09-791-537-72548 (1-772)

QY 200 ATGGCTGTTAACTAGAGATTCTCTTTCTGGGAGTATTACTTACAGCAGACCA 259
Db 1 MetGlyTrpPheThrGlyLeuAlaCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20

QY 260 AACTATCAGATGGGAAGCAACATGTGCCAGGCTGAAATATCTCAAGAAAGATGTG 319
Db 21 AsnTyrAlaAsnGlyLysAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40

QY 320 GAATCCAAATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCAGTTATCATACCTTC 379
Db 41 GluSerAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60

QY 380 CTTTGGATGAGGAACGAGTAGGCTGTATGTTGGAGCAAGGATCAGATATTTTCATTC 439
Db 61 LeuLeuAspGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80

QY 440 GACCTGGTTATATCAAGATTCTCAAAGATTGTGGCCAGTATCTTACACAGAGA 499
Db 81 AsnLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100

QY 500 GATGAATCAAGTGGGCTGGAAAGACATCTCGAAAGAAATGTGCTAAATTTTCAGAGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120

QY 560 CTTAAGGCATATATCAGACTCAGTGTGACGCTGTGGAACGGGGCTTTTCATCCAAAT 619
Db 121 LeuGluAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140

QY 620 TGCACCTACATTAATGGAATCTCATCTGAGGCAATATTTTAAAGCTGGAGCACTCA 679
Db 141 CysThrTyrIleGluValGlyHisHisProGluAspAsnIlePheLysLeuGlnAspSer 160

QY 680 CATTTTGAAACGGCGCTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
```

```
QY 740 TTAATAGATGAGAAATTATCTCTGGAACCTGACAGTGAATTTTATGGGCGAGACTTTGCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaIleAspPheMetGlyArgAspPheAla 200

QY 800 ATCTTCCGAACTCTTGGGCGACCCACCACCAATCAGGACAGACAGCATGATTTCCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220

QY 860 CTCAATGATCAAAGTTCATTAGTGGCCACCTCATCTCAGAGAGTGACAATCTCTGAAGAT 919
Db 221 LeuAsnAspProArgPheIleSerAlaHisLeuIleProGluSerAspAsnProGluAsp 240

QY 920 GACAAAGTATATCTTTCTTCCGTAATGCAATAGATGAGAGCACTCTGGAAAAGCT 979
Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260

QY 980 ACTCACGCTAGATAGTCAAGATATGCAAGATGACTTTTGGAGGCGCAGAAAGTCTGTGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280

QY 1040 AATAATGGACAACATTTCTCAAAGCTCGTCTGATTTTCTCAGTCCAGGTCGCAATGGC 1099
Db 281 AsnLysTrpThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300

QY 1100 ATTGACACTCATTTTGTGATGAACCTGCAGAGATGATTTCTTAATGAATTTTAAAGATCCTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnSerLysAspProLys 320

QY 1160 AATCCAGTTGTATGAGTGTGTTTACGACTTCCAGTAACTTTTCAAGGGATCAGCCGTG 1219
Db 321 AsnProIleValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340

QY 1220 TGATGTATAGCATGAGTGTGAGAGAGGTGTTCTTGTGTCATATGCCACAGGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360

QY 1280 GGACCCAACTATCAATGGTGCCTTATCAAGGAGAGTCCCTTATCCAGCGCCAGGAAT 1339
Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380

QY 1340 TGTCCCAAGCAAAACATTTGGTGGTGTGACTCTCAAGAGACCTTCTCTGATGATGTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400

QY 1400 ACCTTCCAGAGAGTCAATCCAGCATGTACATCCAGTGTTCCTATGAACAATCGCCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProIleAsnAsnArgPro 420

QY 1460 ATAGTGATCAAAACCGAGTGTAAATTTACAAATTTACAAATTTGTCGTAGACCGAGTGGAT 1519
Db 421 IleMetIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440

QY 1520 GCAGAAGATGCACAGTATGATGTTATGTTTATCGGAACAGATGTTGGACCGTCTTTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460

QY 1580 GTAGTTTCAATTCCTAAGAGAGACTTGTGTATGATTTAGAGAGGTTCTGCTGGAGAAATG 1639
Db 461 ValValSerValProLysGluThrTrpHisAspLeuGluValLeuLeuGluGluMet 480

QY 1640 ACAGTTTTCGGGAACCGACTGCTTATTTTCAGCAATGAGCTTTTCCACTTAAGCAGCAACA 1699
Db 481 ThrValPheArgGluProThrThrIleSerAlaMetGluLeuSerThrLysGlnGln 500

QY 1700 CTATATATGTTCAACGGCTGGGTTGCCAGCTCCCTTTACACCGGTGTGATTTTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520

QY 1760 GGGAAACGGTGTGCTGAGTGTGCTGCTCCCGAGACCTTACTGCTGCTGGGATGGTTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
```


1220 TGTATGATATAGCATGATGTGAGAAAGGGTCTCTTGGTCCATATGCCCCACAGGAT 1279
Db
341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
1280 GGACCAACTATCAATGGGTGCTTATCAAGGAAGTCCCTATCCAGGCCAGAACT 1339
Db
361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
1340 TGTCCTCCAGCAAAACATTTGGTGGTTTGAAGTCTACAAAGGACCTTCCGTGATGATCTATA 1399
Db
381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspValIle 400
1400 ACCTTTGCAAGAAGTCCATCCAGCCATGTACAATCCAGTGTCTTCCATGAACAATCGCCCA 1459
Db
401 ThrPheGlyArgSerHisProAlaMetTyrAsnProValPheProIleAsnAsnArgPro 420
1460 ATAGTGATCAAAACCGATTAATATCAATTTACAAATTTGCTGTAGACCGAGTGGAT 1519
Db
421 IleMetIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
1520 GCAGAGATGGACATGATGATCTTATGTTTATCGGACAGAGATGTTGGACCGTCTCTTAAA 1579
Db
441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
1580 GTAGTTTCAATTCCTAAGAGACTTGGTATGATTTAGAGAGGTTCTGTGGAAGAAATG 1639
Db
461 ValValSerValProLysGluThrTrpHisAspLeuGluGluValLeuLeuGluMet 480
1640 ACAGTTTTCGGGAACCGACTGCTATTTCCAGCAATGGAGCTTCCCAAGCAGCAACAA 1699
Db
481 ThrValPheArgGluProThrThrIleSerAlaMetGluLeuSerThrLysGlnGln 500
1700 CTATATATGTTTCAACGCTGGGTGCGGAGTCCCTTTACAGCGGTGATATTTAC 1759
Db
501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
1760 GCGAAGCGGTGCTGAGTGTTCCTCGCCGAGACCCCTTACTGTCTGGATGGTCT 1819
Db
521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
1820 GCATGTTCTCGTATTTTCCCACTGCAAGAGACGACAGCAGCAAGATATAAGAAAT 1879
Db
541 SerCysSerArgTyrPheProThrAlaLysArgArgThrArgGlnAspIleArgAsn 560
1880 GGAGACCACTGACTGCTTACAGCTTACAC---CATGATAATCAACATGCCACAGC 1936
Db
561 GlyAspProLeuThrHisCysSerAspLeuGluAspHisAspAsnHisGlyProSer 580
1937 CCTGAAGAGAGAAATCATCTATGCTGTAGAGAAATAGTACACATTTTGGAAATGCACTCG 1996
Db
581 LeuGluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerPro 600
1997 AAGTCGACAGAGCGCTGCTTATTGGCAATTCAGAGCGCAATGAAGAGCGAAAGAA 2056
Db
601 LysSerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnArgSerLysArg 620
2057 GAGTCAGATGGATGATCATATCATCAGACAGATCAAGSCCTTCTGCTAGTGTCTA 2116
Db
621 GluIleArgMetGlyAspHisIleIleArgThrGluGlnGlyLeuLeuLeuArgSerLeu 640
2117 CAACAGAAGGATTCAGGCAATTTACCTCTGCCATGGGTGGAACATGGGTTTCATACAACT 2176
Db
641 GlnLysLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheMetGlnThr 660
2177 TTCTTTAAGGTAACCTGGAAGTCATGTACACAGAGCAATTTGGAAGAACTTCTTCATAAA 2236
Db
661 LeuLeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLys 680
2237 GATGATGATGGATGGCTCTAAGACCAAGAAATGCTCAATAGCATGACACCTAGCCAG 2296
Db
681 AspAspAspGlyAspGlySerLysIleLysGluMetSerSerSerMetThrProSerGln 700
2297 AAGGTCTGTACAGAGACTTCATGAGCTTCATCAACCCCACTCATCAACAGTGGAT 2356

701 LysValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAsp 720
QY
2357 CAGTCTCTGTGAAAGTTTGAAGAGCGACCGAAACCAACCTCGCAAGCCAGGACAT 2416
Db
721 GluPheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHis 740
QY
2417 ACCCCAGGGAACACTTAACAAATGAAGCACCTTACAGAAAAATAAGAAAGGTAGAAACAGG 2476
Db
741 SerGlnGlySerSerAsnLysTrpLysHisMetGlnGluSerLysLysGlyArgAsnArg 760
QY
2477 AGGACCCACGAATTTGAGAGCGGCCCCAGGAGTGTCT 2512
Db
761 ArgThrHisGluPheGluArgAlaProArgSerVal 772

RESULT 15
US-09-791-537-128280
; Sequence 128280, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 128280
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-128280

Alignment Scores:
Pred. No.: 0 Length: 772
Score: 3989,50 Matches: 730
Percent Similarity: 97,28% Conservative: 21
Best Local Similarity: 94,56% Mismatches: 20
Query Match: 81,48% Indels: 1
DB: 22 Gaps: 1

US-09-774-490-1 (1-2709) x US-09-791-537-128280 (1-772)

QY 200 ATGGCTGGTTAACTAGGATTTGTCTTTCTGGGAGTATTACTTACAGCAGACA 259
Db 1 MetGlyTrpPheThrGlyIleAlaCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGAATGGAGAACAAATGTGCAAGGCTGAAATTTATCTCAAAAGAAATGTTG 319
Db 21 AsnTyrAlaAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAACAAATGTGATCACTTTCAATGGCTTGCCCAACAGCTCCAGTTATCATACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyrHisThrPhe 60
QY 380 CTTTGTGATGAGGAACCGAGTAGCTGTATGTTGGAGCAAGGATCATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACTGGTTAATATACAGGATTTTCAAAGATTTGTGTGGCCAGTATCTTACACCAAGAGA 499
Db 81 AsnLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATGAATCAAGTGGCTGGAAAAACATCTCTGAAGAAATGTCTTAATTTTCATCAAGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATATCAGACTCATTGTACCGCTGTGAAACGGGGCTTTTCATCAATT 619
Db 121 LeuGluAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140

620 TGCACCTACATTGAATTCGGACATCATCTCTGAGGACAATATATTTTAAGTCGAGAACTCA 679
Db CysThrTyrIleGluValGlyHisHisProGluAspAsnIlePheLysLeuGlnAspSer 160
680 CATTTTGAAACGGCGTGGGAAGAGTCCATATGACCCCTAAGCTCCTGACAGCATCCCTT 739
Db HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
740 TTAATAGATGGAGAATTACTCTGGAACCTGCAGCTGATTTTATGGGGGAGACTTTGCT 799
Db LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
800 ATCTTCCGAACTCTGGGCACACACACCAATCAGACAGACAGATGATTCAGGTGG 859
Db IlePheArgThrLeuGlyAspHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
860 CTCATATGATCAAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGACAATCCTCAAGAT 919
Db LeuAsnAspProArgPheIleSerAlaHisLeuIleProGluSerAspAsnProGluAsp 240
920 GACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGGAGAACACTCTGGAAGCT 979
Db AsnLysValTyrPhePhePheArgGluAsnAlaIleGlyGlyGluHisSerGlyLysAla 260
980 ACTCAGCTAGATAGTTCAGATATGCAAGAAATGATCTTTGGAGGGCAGAGATCTGTG 1039
Db ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
1040 AATAATGACAACTCTCAAGCTCGTCTGATTTCTGATTCAGTCCAGTCCCAATGCG 1099
Db AsnLysIleTyrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
1100 ATTGACACTCAATTTTGATGAATCGCAGGATGTATTCCTAATGAACTTTAAAGATCCTAAA 1159
Db IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnSerLysAspProLys 320
1160 AATCAGTGTATATGGAGTGTATAGACTTCCAGTCCAGTAACTTTCAAGGGATCAGCGTG 1219
Db AsnProIleValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
1220 TGTATGTATAGCATGATGATGAGAGGGTGTCTTGTGTCATATGCCACAGGAT 1279
Db CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
1280 GGACCCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTATCCAGCGCCAGAACT 1339
Db GlyProAsnTyrGlnTyrValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
1340 TGTCACAGCAAAACATTTGGTGGTGTGACTCTACAAGGACCTTCTCGATGATGTTATA 1399
Db CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
1400 ACCTTTGCAAGAGTCAATCCAGCATGACATCCAGTGTTCCTATGAACTCGCCCA 1459
Db ThrPheGlyArgSerHisProAlaMetTyrAsnProValPheProIleAsnAsnArgPro 420
1460 ATAGTGATCAAAACGGATGTAATTTATCAAAATTTGCTGAGACCGAGTGAT 1519
Db IleMetIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
1520 CGAAGATGACAGTATGATGTTATGTTATCGGAACAGATGTTGGGACCGTCTCTAAA 1579
Db AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
1580 GTAGTTTCAATCTCAAGGACCTGTGATGTTTATAGAGAGTTCCTCTGGAAGAAATG 1639
Db ValValSerValProLysGluThrTrpHisAspLeuGluValLeuLeuGluMet 480
1640 ACAGTTTTTCCGGAACCCAGCTCTATTTCAAGAAATGGAGCTTTCACATAAGCAGCAACA 1699
Db ThrValPheArgGluProThrThrIleSerAlaMetGluLeuSerThrLysGlnGln 500
1700 CTATATATTTGGTTCAACGGCTGGGTTGCCAGCTCCCTTTTACACCGGTGTGATTTTAC 1759

501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
1760 GGGAAAGCGTGTGCTGAGTGTTCCTCCCGAGACCTTACTGTGCTTGGATGGTCTCT 1819
Db GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
1820 GCATGTTCTCCTATTTTCCCACTGCCAAGAGACGCACAAGACGACGACGATATAAGAAAT 1879
Db SerCysSerArgTyrPheProThrAlaLysArgArgThrArgGlnAspIleArgAsn 560
1880 GGAGACCCACTGACTCCTGCTTCAGACTTACAC---CATGATAATCACCATGGCCACAGC 1936
Db GlyAspProLeuThrHisCysSerAspLeuGluAspHisAspAsnHisGlyGlyProSer 580
1937 CCTGAAGAGAGAATCATCTATGGTGTAGAGAAATAGTAGACATTTTTCGAATCGAGTCCG 1996
Db LeuGluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerPro 600
1997 AAGTCGACAGAGCGCTGCTCTATTGGCAATTCAGAGCGCAAAATGAAGAGCGAAAGAA 2056
Db LysSerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnArgSerLysArg 620
2057 GAGATCAGATGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGCTAGTCTA 2116
Db GluIleArgMetGlyAspHisIleIleArgThrGluGlnGlyLeuLeuLeuArgSerLeu 640
2117 CAACAGAGAGGATTCAGGCAATTTACCTCTGCCATCGGTGGGAACATGGGTTCATACAACT 2176
Db GlnLysLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheMetGlnThr 660
2177 CTTCTTAAGGTAACCCCTGGAAGTCTATTCACACAGACGATTTTGGAAAGAACTTCTTCAAAA 2236
Db LeuLeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuHisLys 680
2237 GATGATGATGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAG 2296
Db AspAspAspGlyAspGlySerLysIleLysGluMetSerSerSerMetThrProSerGln 700
2297 AAGGTCTGTACAGAGACTTCTATGCAGTCTATCAACACCCCACTCTCAACACGATGGAT 2356
Db LysValTyrTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAsp 720
2357 GAGTTCTGTGAACAAGTTTGGAAAAGGACCGCAAAACAACTCGGCAAGGCCAGGACAT 2416
Db GluPheCysGluGlnValTyrLysArgAspArgLysGlnArgGlnArgProGlyHis 740
2417 ACCCGGGAACAGTAACAATGGAAGCATTACAGAAATATAGAAAGGTAGAAACAGG 2476
Db SerGlnGlySerSerAsnLysTrpLysHisMetGlnGluSerLysLysGlyArgAsnArg 760
2477 AGGACCCACGAAATTTGAGAGGGCCACCCAGGAGTGTTC 2512
Db ArgThrHisGluPheGluArgAlaProArgSerVal 772

Search completed: August 3, 2003, 10:19:46
Job time : 462 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 10:01:29 ; Search time 12 Seconds
(without alignments)
5659.545 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 4896

Sequence: 1 aatctttattttatcgatg.....agggttttttctcctaacc 2709

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 67526 seqs, 12534983 residues

Total number of hits satisfying chosen parameters: 135052

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-O=/cgn2_1/USPTO_epool/US09774490/runat_03082003_095417_25391/app_query.fasta_1.2887
-DB=Pending Patents AA New -QFMT=fastan -SUFFIX=rapn -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOLALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09774490 @CGN 1.1 @runat_03082003_095417_25391 -NCPU=6 -ICPU=3
-NO_MWAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Pending Patents AA New:*
2: /cgn2_6/ptodata/2/paa/pct NEW COMB.pcp:*
3: /cgn2_6/ptodata/2/paa/US06 NEW COMB.pcp:*
4: /cgn2_6/ptodata/2/paa/US07 NEW COMB.pcp:*
5: /cgn2_6/ptodata/2/paa/US08 NEW COMB.pcp:*
6: /cgn2_6/ptodata/2/paa/US09 NEW COMB.pcp:*
7: /cgn2_6/ptodata/2/paa/US10 NEW COMB.pcp:*
8: /cgn2_6/ptodata/2/paa/US60 NEW COMB.pcp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2044	41.7	775	1	PCT-US03-09929-92
2	2044	41.7	775	1	PCT-US03-09929-94
3	2044	41.7	775	1	PCT-US03-09929-96
4	2044	41.7	775	1	PCT-US03-09929-98
5	1917	39.2	636	6	US-10-408-765A-2003
6	1302.5	26.6	514	6	US-10-357-820-28
7	1136.5	23.2	431	6	US-10-369-072-97
8	966	19.7	834	6	US-10-369-072-47
9	964	19.7	430	6	US-10-369-072-91
10	952.5	19.5	638	1	PCT-US03-09929-10
11	952.5	19.5	833	6	US-10-369-072-14
					Sequence 92, Appl
					Sequence 94, Appl
					Sequence 96, Appl
					Sequence 98, Appl
					Sequence 2003, Ap
					Sequence 28, Appl
					Sequence 47, Appl
					Sequence 91, Appl
					Sequence 10, Appl
					Sequence 14, Appl

12	950.503	19.4	833	1	PCT-US03-09929-6
13	950.5	19.4	833	6	US-10-369-072-16
14	950.5	19.4	833	6	US-10-369-072-18
15	949.503	19.4	634	1	PCT-US03-09929-2
16	949.5	19.4	638	1	PCT-US03-09929-8
17	949.503	19.4	833	1	PCT-US03-09929-12
18	943	19.3	963	6	US-10-369-072-46
19	937	19.1	838	6	US-10-369-072-49
20	889	18.2	893	6	US-10-369-072-38
21	857.5	17.5	782	6	US-10-369-072-39
22	856.503	17.5	666	1	PCT-US03-09929-56
23	855.5	17.5	878	1	PCT-US03-09929-54
24	854.503	17.5	939	1	PCT-US03-09929-32
25	854	17.4	640	1	PCT-US03-09929-16
26	854	17.4	649	1	PCT-US03-09929-50
27	854	17.4	861	1	PCT-US03-09929-52
28	854	17.4	1035	1	PCT-US03-09929-18
29	854.003	17.4	1047	1	PCT-US03-09929-14
30	854.003	17.4	1047	1	PCT-US03-09929-48
31	852.503	17.4	626	1	PCT-US03-09929-34
32	852.5	17.4	1018	1	PCT-US03-09929-28
33	850.503	17.4	998	1	PCT-US03-09929-20
34	839.5	17.1	981	1	PCT-US03-09929-30
35	834	17.0	756	6	US-10-369-072-8
36	827.5	16.9	1017	6	US-10-451-010-2
37	825.5	16.9	754	6	US-10-369-072-10
38	825.5	16.9	761	6	US-10-293-244-3947
39	825.003	16.9	495	1	PCT-US03-09929-4
40	817	16.7	1070	6	US-10-461-791-8
41	817	16.7	1070	6	US-10-461-792-8
42	817	16.7	1086	6	US-10-461-791-4
43	817	16.7	1086	6	US-10-461-792-4
44	816.5	16.7	641	6	US-10-461-791-13
45	816.5	16.7	641	6	US-10-461-792-13

ALIGNMENTS

RESULT 1

PCT-US03-09929-92
; Sequence 92, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: Curaseq version 0.1
; SEQ ID NO 92
; LENGTH: 775
; TYPE: PRT

ORGANISM: Homo sapiens
PCT-US03-09929-92

Alignment Scores:

Pred. No.: 7,64e-165 Length: 775
Score: 2044.00 Matches: 374
Percent Similarity: 67.57% Conservative: 147
Best Local Similarity: 48.51% Mismatches: 234
Query Match: 41.75% Indels: 16
DB: 1 Gaps: 8

US-09-774-490-1 (1-2709) x PCT-US03-09929-92 (1-775)

```
QY 218 ATTGCTGCTCTTTCTGGGAGTATTACTTACAGCAGAGCAAACTATCATCAAGTGGGAG 277
DB 8 IIEThrLeuLeuLeuTrpGlyTyrLeuLeuGluLeuTrpThrGlyGlyHisThrAlaAsp 27
QY 278 AACAAATGTCAGGCTGAAATATCTACAAAGAAATGTTGGAATCCAAATGTGATC 337
DB 28 ThrThrHisProArgLeuArgLeuSerHisLysGluLeuLeuAsnLeuAsnArgThrSer 47
QY 338 ACTTTCATGGCTGGCCAAACAGCTCCAGTTATCATCTCTCTTTGGATGAGAACGG 397
DB 48 IIEPheHisSerProPheGlyPheLeuAspLeuHisThrMetLeuLeuAspGluTyrGln 67
QY 398 AGTAGGCTGTATGTTGGACCAAGGATCACATATTTTTCATTCGACTGCTGTTAATCAAG 457
DB 68 GluArgLeuPheValGlyGlyArgAspLeuValTyrSerLeuSerLeuGluArgIleSer 87
QY 458 GAT---TTTCAAAGATTGTGGCCAGTATCTTACACCAAGAGAGATGAATGCAAGTGG 514
DB 88 AspGlyTyrLysGluIleHisTrpProSerThrAlaLeuLysMetGluGluCysIleMet 107
QY 515 GCTGGAAGAAGACATCCTGAAAGATGTGCTAATTTTCATCAAGGTTACTTAAAGCATATAT 574
DB 108 LysGlyLysAsp---AlaGlyGluCysAlaAsnTyrValArgValLeuHisHisTyrAsn 126
QY 575 CAGACTCACTGTAGCCCTGTGGACGGGGCTTTTTCATCCAAATTTGACCTTACATTCGAA 634
DB 127 ArgThrHisLeuLeuThrCysGlyThrGlyAlaPheAspProValCysAlaPheIleArg 146
QY 635 ATTGACATCATCTCGAGACAAATATTTTAACTGGAGAACTCATCTTTGAAACGGC 694
DB 147 ValGlyTyrHisLeuGluAspProLeuPheHisLeuGluSerProArgSerGluArgGly 166
QY 695 CTGGGAAGAGTCATATGACCTTAAGCTGTGCTGACAGCATCCCTTTTAAATAGATGGAGAA 754
DB 167 ArgGlyArgCysProPheAspProSerSerSerPheIleSerThrLeuIleGlySerGlu 186
QY 755 TTATCTCTGGAAGTGCAGCTCATTTTATGGGCGAGACTTTTGTCTATCTTCGGAACCTTT 814
DB 187 LeuPheAlaGlyLeuTyrSerAspTyrTrpSerArgAspAlaAlaIlePheArgSerMet 206
QY 815 GGGCACCACCAACCAATCAGGACAGACAGCATGATTCAGGTGCTCAATGATCCAAAG 874
DB 207 GlyArgLeuAlaHisIleArgThrGluHisAspAspGluArgLeuLeuLysGluProLys 226
QY 875 TTCATTAGTCCCACTCATCTCAGAGAGTGACATCTCTGAAGATGACAAAGTATACCTTT 934
DB 227 PheValGlySerTyrMetIleProAspAsnGluAspArgAspAsnLysValTyrPhe 246
QY 935 TTCTTCCGTAAGTGCATAGATGGAGAACACTCTGGAAGAGTCTACGCTAGAATA 994
DB 247 PhePheThrGluLysAlaLeuGluAlaGluAsnAlaHisAlaIleTyrThrArgVal 266
QY 995 GGTCAATATGCAAGAAATGACTTTGGAGGGCAGAGAGTCTGGTGAATAAATGGCAACA 1054
DB 267 GlyArgLeuCysValAsnAspValGlyGlyGlnArgIleLeuValAsnLysTrpSerThr 286
QY 1055 TTCTTCAAAGCTCGTCTGATTTGCTCAGTCCAGGTCCTCAATGCGATGACACTCATTTT 1114
DB 287 PheLeuLysAlaArgLeuValCysSerValProGlyMetAsnGlyIleAspThrTyrPhe 306
```

```
QY 1115 GATGAATGCTCAGGATGATTCTCTAATGAATCTTTAAAGATCCTAAATCCAGTTGTATAT 1174
DB 307 AspGluLeuGluAspValPheLeuLeuProThrArgAspHisLysAsnProValIlePhe 326
QY 1175 GGAGTGTATTACGACTTCCAGTAACATTTTCAAGGATCAGCCGTGTGTATGTATAGCATG 1234
DB 327 GlyLeuPheAsnThrThrSerAsnIlePheArgGlyHisAlaIleCysValTyrHisMet 346
QY 1235 AGTCATGTGAGAAGGGTGTCTTGTGTCATATCCACAGGATGGACCCCAACTATCAA 1294
DB 347 SerSerIleArgAlaAlaPheAsnGlyProTyrAlaHisLysGluGlyProGluTyrHis 366
QY 1295 TGGGTGCTTATCAAGGAAGAGTCCCTATCCAGCGCCAGCAACTGTGTCCAGCAAAACA 1354
DB 367 TrpSerValTyrGluGlyLysValProTyrProArgProGlySerCysAlaSerLysVal 386
QY 1355 TTTGCGTGT---TTTGACTCTACAAGACCTTCTCTGATGATGTATTAACCTTTGCAAGA 1411
DB 387 AsnGlyGlyArgTyrGlyThrThrLysAspTyrProAspAspAlaIleArgPheAlaArg 406
QY 1412 AGTCATCCAGCCATGATCAATCCAGTGTTCCTATGACATCGCCCAATAGTGATCAA 1471
DB 407 SerHisProLeuMetTyrGlnAlaIleLysProAlaHisLysLysProIleuValLys 426
QY 1472 ACGGATGTAAATTAATCAATTTTACAAATTTCTGCTAGACCGAGTGGATGCAAGATGGA 1531
DB 427 ThrAspGlyLysTyrAsnLeuLysGlnIleAlaValAspArgValGluAlaGluAspGly 446
QY 1532 CAGTATGATGTATGTTTATCGGAACAGATGTTGGACCGTCTTAAAGTAGTTTCAATT 1591
DB 447 GlnTyrAspValLeuPheIleGlyThrAspAsnGlyIleValLeuLysValIleThrIle 466
QY 1592 CCTAAGGAGACTTGGTATGATTTAGACAGCTTCTGCTGGAGAAATACACAGTTTTCGG 1651
DB 467 TyrAsnGlnGluMetGluSerMetGluGluValIleLeuGluGluLeuGluIlePheLys 486
QY 1652 GAACCGAGCTGATTTTTCAGCAATCGAGCTTTCCACTAAGCAACCAACTATATATGTT 1711
DB 487 AspProValProIleIleSerMetGluIleSerSerLysArgGlnGlnLeuTyrIleGly 506
QY 1712 TCAACGCTGGGTGGTCCAGCTCCCTTTACCGGTGTGTATTTTACCGGAAGACCGTGT 1771
DB 507 SerAlaSerAlaValAlaGlnValArgPheHisHisCysAspMetTyrGlySerAlaCys 526
QY 1772 GCTCAGTGTGCTCCGCGGAGACCTTACTGTCTTGGGATGTTCTGTCATGTTCTCGC 1831
DB 527 AlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyIleSerCysSerArg 546
QY 1832 TATTTTCCCACT-----GCAAAGAGACGCAACAGACACAAGATATTAAGAAATCGA 1882
DB 547 TyrTyrProThrGlyThrHisAlaLysArgArgPheArgArgGlnAspValArgHisGly 566
QY 1883 GACCACTGACTACTGTTTCAGACTTACCATGATATTCACCATGTCACCATGGCCAGCCCTGA 1942
DB 567 AsnAlaAlaGlnGlnCysPheGlyGlnGlnPheValGlyAspAlaLeuAspLysThrGlu 586
QY 1943 GAGAGATCATCTATGTTGTAGACATATTTTGAATGACATTTTGGATCCAGTCCGAAAGTCG 2002
DB 587 GluHisLeuAlaTyrGlyIleGluAsnAsnSerThrLeuLeuGluCysThrProArgSer 606
QY 2003 CAGAGAGCGCTGTCTATTGGCAATTCAGAGGCGGAAATGAAGAGCGGAAAGAGATC 2062
DB 607 LeuGlnAlaLysValIleTrpPheValGlnLysGlyArgGluThrArgLysGluGluVal 626
QY 2063 AGATGTGATGATCATCATCAGGACAGATCAAGCCTTCTGCTACGTAGTCTACAAACAG 2122
DB 627 LysThrAspAspArgValValLysMetAspLeuGlyLeuLeuPheLeuArgLeuHisLys 646
QY 2123 AAGGATTCAGCAATTAACCTCTGTCATCGGTGGAACATGGGTTTCATCAACAACCTCTTCT 2182
DB 647 SerAspAlaGlyThrTyrPheCysGlnThrValGluHisSerPheValHisThrValArg 666
QY 2183 AAGGTAACCCCTGGGAAGTCATTTGACACAGACATTTTGAAGAACTTCTTTCATAAAGATGAT 2242
```


Qy	1355	TTTGGTGGT---TTTGGACTCTACAAAGGACCTTCTCTGATGATGTATAAACCTTTTGCAAGA	1411
Db	387	AsnGlyGlyArgTyGlyThrThyLysAspTyrProAspAspAlaIleArgPheAlaArg	406
Qy	1412	AGTCATCCAGCCATGTACATCCAGTGTTCCTATGAACAATCCCCCAATAGTGTACAA	1471
Db	407	SerHisProLeuMetTyrGlnAlaIleLysProAlaHisLysLysProIleLeuValLys	426
Qy	1472	ACGGATGTAATATTCAATTTACACAATTCCTAGACCGAGTGGATGAGAAAGATGGGA	1531
Db	427	ThrAspGlyLysTyrAsnLeuLysGlnIleAlaValAspArgValGlnAlaGluAspGly	446
Qy	1532	CAGTATGATGTTATGTTTATCCGAACACAGATGTTCCGACCGCTTCTTAAGTAGTTCAT	1591
Db	447	GlnTyrAspValLeuPheIleGlyThrAspAsnGlyIleValLeuLysValIleThrIle	466
Qy	1592	CCTAAGGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGAGAAATGACAGTTTTCGG	1651
Db	467	TyrAsnGlnGluMetGluSerMetGlnGluValIleLeuGluGluLeuGlnIlePheLys	486
Qy	1652	GAACCGACTGCTATTCTCAGCAATCGAGCTTTCACCTAAGCAGCAACAATATATATGGT	1711
Db	487	AspProValProIleIleSerMetGlnIleSerSerLysArgGlnGlnLeuTyrIleGly	506
Qy	1712	TCAACGGCTGGGTGGCCAGCTCCCTTTACACCGGTGTGATATTTACGGGAAGCGTGT	1771
Db	507	SerAlaSerAlaValAlaGlnValArgPheHisLysCysAspMetTyrGlySerAlaCys	526
Qy	1772	GCTGAGTGTGGCTCGCCGAGACCCCTACTGCTGGATGCTGGATGCTGCTGCTCTCCG	1831
Db	527	AlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyIleSerCysSerArg	546
Qy	1832	TATTTTCCCACT-----GCAAGAGAGCGCAACAGACACAAGATATAAGAAATGGGA	1882
Db	547	TyrTyrProThrGlyThrHisAlaLysArgArgPheArgGlnAspValArgHisGly	566
Qy	1883	GACCCACTGACTACTGTTTCAGACTTACCATCATGATNAATCACCATGGCCACAGCCCTGAA	1942
Db	567	AsnAlaAlaGlnGlnCysPheGlyGlnGlnPheValGlyAspAlaLeuAspLysThrGlu	586
Qy	1943	GAGAGAATCATCTATGTTGTAGAGAAATAGTAGACATTTTGGAAATCGCTCGGAAGTCG	2002
Db	587	GluHisLeuAlaTyrGlyIleGluAsnAsnSerThrLeuLeuGluCysThrProArgSer	606
Qy	2003	CAGAGCGCTGGTCTATTGGCAATCCAGAGCGGAAATGAAGAGCGAAAGAGAGATC	2062
Db	607	LeuGlnAlaLysValIleTyrPheValGlnLysGlyArgGlnThrArgLysGluGluVal	626
Qy	2063	AGAGTGGATGATCATATCATCAGACACAGATCAAGGCCTTCTGCTACGTCTTACAAAC	2122
Db	627	LysThrAspAspArgValValValLysMetAspLeuGlyLeuLeuPheLeuArgLeuHisLys	646
Qy	2123	AAGGATTACAGCAATTAACCTCTGCGATCGGTGGAAACATGGTTTTCATAAACTTCTT	2182
Db	647	SerAspAlaGlyThrTyrPheCysGlnThrValGluHisSerPheValHisThrValArg	666
Qy	2183	AAGGTAACCCCTGGGAAGTCATTGACACAGACGATTTGGAAGAACTTCTTCAATAAGATGAT	2242
Db	667	LysIleThrLeuGluValValGluGluGluLysValGluAspMetPheAsnLysAspAsp	686
Qy	2243	GATGAGATGGCTCTAAGACCAA-----GAAATGTCCATAGCATGACACCTAGCCAG	2296
Db	687	GluGluAspArgHisLysArgMetProCysProAlaGlnSerSerIleSerGlnGlyAla	706
Qy	2297	AAGGTCTGGTACAGAGACTTCATCGACTCATCAACCAACCCCAATCTCAACACGATGGAT	2356
Db	707	LysProIleTyrLysGluPheLeuGlnIleGlyTyrSerAsnPheGlnArgValGlu	726
Qy	2357	GAGTTCGTGAAACAGTTTGAAGGACCGAACAACATGCTCGGCAAGGCCACGACAT	2416
Db	727	GlnTyrCysGluLysValTyrCysThrAspArgLysArgLysLeuLysMet-----	744

```

QY      2417 ACCCCAGGGAACAGTAAACAATGGAGCACTTACAA-----GAAATAAGAAAAGTGTAACAAC 2476
Db      ::::|||||:|||||:::   |||||   |||:::
       745 -----SerProSerLysTrpLysTyraAsnProGlnGluLysIlyLeuArgSer 761
QY      2474 AGGAGGACCACCAATTGTGAGAGGCCACCCAGG 2506
Db      :::   |||   |||   |||||
       762 LysProGluHis-----TyrArgLeuProArg 770

RESULT 3
PCT-US03-09929-96
; Sequence 96, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; REMAINING PRIOR APPLICATION DATA REMOVED - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuroSeqdist version 0.1
; SEQ ID NO 96
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-09929-96

Alignment Scores:
Pred. No.:          7.64e-165           Length:          775
Score:             2044.00              Matches:         374
Percent Similarity: 67.57%              Conservative:    147
Best Local Similarity: 48.51%            Mismatches:     234
Query Match:        41.75%               Indels:         16
DB:                1                    Gaps:          8

US-09-774-490-1 (1-2709) x PCT-US03-09929-96 (1-775)
QY      218 ATTCTCTGTCCTTTCTGGGAGATTACTTACAGCAGCAAGCAAACCTATCAGAATGGGAAG 277
Db      :|||   |||   |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
       8 IleThrLeuLeuLeuTrpGlyTyrLeuLeuGluLeuTrpThrGlyGlyHisThrAlaAsp 27
QY      278 AACAAATGTGCACAGCGTGAATATCTCTCAAAAGAAATGTGTGAATCCAACAATGTGTATC 337
Db      :|||   |||   |||   |||   |||   |||   |||   |||   |||   |||   |||   |||
       28 ThrThrHisProArgLeuArgLeuSerHisLysGluLeuLeuAsnLeuAsnArgThrSer 47
QY      338 ACTTTCAAATGGCTTGGCCAACACAGCTCAGTTATCATACCTTCCTTTTTGGATGAGGAACGG 397
Db      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
       48 IlePheHisSerProPheGlyPheLeuAspLeuHisThrMetLeuLeuAspGluTyrGln 67
QY      398 AGTAGGCTGTATGTGGAGCAAAAGGATCACATATTTTCATTTCGACCTGGTTAATATCAAG 457
Db      :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||   :|||
       68 GluArgLeuValGlyGlyArgGlyValTyrSerLeuSerLeuGluArgIleSer 87
QY      458 GAT----TTTCAAAGAGTTGTGTGGCCAGTATCTTACACCAAGAGAGATGAATGCAAGTGG 514

```



```

Db 88 AspGlyTyrIleGluHisIleProSerThrAlaLeuLysMetGluCysIleMet 107
||| ..... ||| ||||| ||||| ..... ||| |||||
QY 515 GCTGGAAGACATCCTGAAAGATGTGCTAAATTTTCATCAAGGTACTTAAGGCATATAAT 574
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 108 LysGlyLysAsp--AlaGlyGluCysAlaAsnTyrValArgValIleHisTyrAsn 126
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 575 CAGACTCACTTGTCAGCCCTGTGGAGCGGGGCTTTTCATCCAAATTTGCACCTACATGAA 634
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 127 ArgThrHisLeuLeuThrCysGlyThrGlyAlaPheAspProValCysAlaPheIleArg 146
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 635 ATTGGACATCATCCTGAGACAAATATTTTAAGCTGGAGAACTCACATTTTGAACACGC 694
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 147 ValGlyTyrHisLeuGluAspProLeuPheHisLeuGluSerProArgSerGluArgGly 166
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 695 COTGGGAAGAGTCCATATGACCTTAAGCTGCTCAGACAGTCCCTTTTATAGATGGAGAA 754
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 167 ArgGlyArgCysProPheAspProSerSerPheIleSerThrLeuIleGlySerGlu 186
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 755 TTATCTCTGGAACCTGACCTGATTTTATGGGCGAGACTTTGCTATCTTCCGAATCTT 814
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 187 LeuPheAlaGlyLeuTyrSerAspTyrTrpSerArgAspAlaAlaIlePheArgSerMet 206
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 815 GGGCACCACCCCAATCAGGACAGCAGCATGATCCAGTGGCTCAATCATCAAG 874
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 207 GlyArgLeuAlaHisIleArgThrGluHisAspAspGluArgLeuLysGluProLys 226
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 875 TTCATTAGTCCACCTCATCTCAGAGATGCAATCCTGAGATGACAAAGTATATCTTT 934
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 227 PheValGlySerTyrMetIleProAspGluAspArgAspAsnLysValTyrPhe 246
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 935 TTCCTCCGTAAGTCAATAGATGGAGAACACTCTGGAAAGTACTCACGCTAGAATA 994
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 247 PhePheThrGluLeuAlaLeuGluAlaGluAsnAlaHisAlaIleTyrThrArgVal 266
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 995 GGTGATATGCAAGAATGACTTTGGAGGACACAGTCTGGTCAATAAATGGACAACA 1054
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 267 GlyArgLeuCysValAsnAspValGlyGlyGlnArgIleLeuValAsnLysTrpSerThr 286
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1055 TTCTCAAAGTCTGCTGATTTGGCTCAGTGCCAGGTCCAAATGGCATTTGACACTATTTT 1114
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 287 PheLeuLysAlaArgLeuValCysSerValProGlyMetAsnGlyIleAspThrTyrPhe 306
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1115 GATGAATCGCAGGATGATTCCTAATGAATCTTAAAGATCTTAAATCCAGTGTATAT 1174
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 307 AspGluLeuGluAspValPheLeuLeuProThrArgAspHisLysAsnProValIlePhe 326
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1175 GGAGTGTTCAGACTTCCAGTAACATTTTCAAGGATCAGCGCTGTGTATAGCATG 1234
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 327 GlyLeuPheAsnThrThrSerAsnIlePheArgGlyHisAlaIleCysValTyrHisMet 346
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1235 AGTGATGTGAGAGGGTGTCTCTGCTCATATGCCAGGATGGACCCCACTATCAA 1294
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 347 SerSerIleArgAlaAlaPheAsnGlyProTyrAlaHisLysGluGlyProGluTyrHis 366
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1295 TGGTGCCCTTATCAAGGAAGTCCCTATCCAGCGCCAGGAACCTGTCAGCAAAACA 1354
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 367 TrpSerValTyrGluGlyLysValProTyrProArgProGlySerCysAlaSerLysVal 386
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1355 TTGTGTGCT--TTTGACTCTCAAGAGACCTTCTGTATGATGTATTAACCTTTCAGA 1411
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 387 AsnGlyGlyArgTyrGlyThrThrLysAspTyrProAspAspAlaIleArgPheAlaArg 406
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1412 AGTCATCCAGCCATGATCAATCCAGTGTCTCTATGAACATCGCCCAATAGTATCAA 1471
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 407 SerHisProLeuMetTyrGlnAlaIleLysProAlaHisLysLysProIleLeuValLys 426
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1472 ACGGATGAATATCAATTTACAAATTTGCTAGACCGAGTGGATGACGAGATGGA 1531
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 427 ThrAspGlyLysTyrAsnLeuLysGlnIleAlaValAspArgValGluAlaGluAspGly 446
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1532 CAGTATGATGTATTTATCGGAACAGATGTTGGACCGTCTTAAAGTATGTTCAATT 1591
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 447 GlnTyrAspValLeuPheIleGlyThrAspAsnGlyIleValLeuLysValIleThrIle 466
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1592 CCTAAGGAGACTTGTATGATTTAGAGAGGTTCCTGCGAAGAAATGACAGCTTTTCGG 1651
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 467 TyrAsnGlnGluMetGluSerMetGluValIleLeuGluLeuGlnIlePheLys 486
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1652 GAACCGACTGCTATTTCCAGCAATGAGCTTTCCACTAAGCAGCAACAATATATATTTGGT 1711
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 487 AspProValProIleSerMetGluIleSerSerLysArgGlnGlnLeuTyrIleGly 506
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1712 TCAACGGCTGGGTGGCCAGCTCCCTTTACACGGTGTATATTTAGCGAAACCGTGT 1771
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 507 SerAlaSerAlaValAlaGlnValArgPheHisHisCysAspMetTyrGlySerAlaCys 526
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1772 GCTGAGTGTTCCTCGCCGAGACCTTACTGTCTGGGATGGTTCGTCATGTTCTCGC 1831
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 527 AlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyIleSerCysSerArg 546
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1832 TATTTTCCACT-----GCAAAGAGACGCAACAGCAGCAAGATATAGAATGGA 1882
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 547 TyrTyrProThrGlyThrHisAlaLysArgArgPheArgGlnAspValArgHisGly 566
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1883 GACCCTGACTCACTGTTCAAGCTTACACCATGATATCACCATGGCCACAGCCCTGAA 1942
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 567 AsnAlaAlaGlnGlnCysPheGlyGlnPheValGlyAspAlaLeuAspLysThrGlu 586
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1943 GAGAAATCATCTATGTTAGAGAAATAGTAGACATTTTGGATGAGTCCGAGATCG 2002
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 587 GluHisLeuAlaTyrGlyIleGluAsnAsnSerThrLeuLeuGluCysThrProArgSer 606
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2003 CAGAGCGCTGGTCTATTGGCAATTCAGAGGCGAAATGAAGACGCAAGAAAGAGATC 2062
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 607 LeuGlnAlaLysValIleTrpPheValGlnLysGlyArgGluThrArgLysGluVal 626
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2063 AGAGTGGATGATCATATCATCAGGACAGATCAAGCCTTCTGCTACGTAGTCTACAACAG 2122
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 627 LysThrAspArgValValLysMetAspLeuGlyLeuLeuPheLeuArgLeuHisLys 646
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2123 AAGGATTCAGGCAATTACTCTGCTGCGTGAACATGGTTCATACAACTCTTCTT 2182
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 647 SerAspAlaGlyThrTyrPheCysGlnThrValGluHisSerPheValHisThrValArg 666
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2183 AAGTAAACCTTGAAGTCAATTCAGACAGCATTTTGGAAAGACTTCTTCATAAAGATCAT 2242
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 667 LysIleThrLeuGluValValGluGluLysValGluAspMetPheAsnLysAspAsp 686
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2243 GATGAGATGCTCTAAGACCAA-----GAAATGTCATAGCATCACACCTAGCCAG 2296
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 687 GluGluAspArgHisHisArgMetProCysProAlaGlnSerSerIleSerGlnGlyAla 706
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2297 AAGTCTGGTACAGAGACTTCATGCGTCTATCAACACCCCAATCTCAACAGATGAT 2356
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 707 LysProTyrTyrLysGluPheLeuGlnLeuIleGlyTyrSerAsnPheGlnArgValGlu 726
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2357 GAGTCTGTGAACAAAGTTTGGAAAAAGGACCGAAAAACACGCTCGCAAGCCAGGACAT 2416
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 727 GluTyrCysGluLysValTyrCysThrAspArgLysArgLysLysLeuLysMet----- 744
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2417 ACCCGGGAACAGTACAAATGAAGCATTACAA-----GAAATGAAGAGGTAGAAAC 2473
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 745 -----SerProSerLysTrpLysTyrAlaAsnProGlnGluLysLysLeuArgSer 761
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 2474 AGGAGGCCACCAAGATTTGAGAGGCCACCCAGG 2506
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 762 LysProGluHis-----TyrArgLeuProArg 770
||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

RESULT 4

PCT-US03-09929-98
 ; Sequence 98, Application PC/TUS0309929
 ; GENERAL INFORMATION:
 ; APPLICANT: Curagen Corporation, et al.
 ; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHODS
 ; FILE REFERENCE: 21402-573B-061

695 CGTGGGAGAGTGCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAAATAGATGGAGAA 754
Db ArgGlyArgCysProPheAspProSerSerPheIleSerThrLeuGlySerGlu 186
755 TTATACTCTGGAACTGCAGCTGATTTTATGGGCGAGACTTCTCTATCTTCCGAACCTCT 814
Db LeuPheAlaGlyLeuTyrSerAspTyrTrpSerArgAspAlaIlePheArgSerMet 206
815 GGGCACCCACCACCAATCAGGACAGACAGCATGATTCAGGTGGCTCAATGATCAAAAG 874
Db GlyArgLeuAlaHisIleArgThrGluHisAspAspGluArgLeuLeuLysGluProLys 236
875 TTCATTAGTCCCACTCATCTCAGAGAGTGACAATCTCTGAAGATGACAAAGATATACTTT 934
Db PheValGlySerTyrMetIleProAspAsnGluAspArgAspAsnLysValTyrPhe 246
935 TTCCTTCGGTCAAAATCCATAGATGGAGACACTCTCGGAAAAGCTACTCACGCTCAAGA 994
Db PhePheThrGlnLysAlaLeuGluAlaGluAsnAsnAlaHisAlaIleTyrThrArgVal 266
995 GGTTCAGATGCAAGAAATGACTTTTGAGGGCGCACAGAAAGTCTGGTGAATAAATGGACAACA 1054
Db GlyArgLeuCysValAsnAspValGlyGlyGlnArgIleLeuValAsnLysTrpSerThr 286
1055 TTCCTCAAGAGTCGTCGTGATTCCTCAGTGCAGTCGCGGTCGCAAAATGGCATTCACATCTATTTT 1114
Db PheLeuLysAlaArgLeuValCysSerValProGlyMetAsnGlyIleAspThrTyrPhe 306
1115 GATGAACTCAGAGATGATTTCTTAATGAATTTTAAAGATCCCTAAAAATCCAGTGTGTAT 1174
Db AspGluLeuGluAspValPheLeuLeuProThrArgAspHisLysAsnProValIlePhe 326
1175 GGAAGTGTTCACGACTTCCAGTAACTTTTCAAGGGATCAGCCGTGTGTATGTATATGATG 1234
Db GlyLeuPheAsnThrThrSerAsnIlePheArgGlyHisAlaIleCysValTyrHisMet 346
1235 AGTGATGTGAGAGGGTGTCTTGTGTCATATGCCCAAGGATGGACCAACTATCAA 1294
Db SerSerIleArgAlaAlaPheAsnGlyProTyrAlaHisLysGluGlyProGluTyrHis 366
1295 TGGGTGCTTATCAAGAGAGAGTCCCTATCCACGCCGAGGAACTGTGCCAGCAAAACA 1354
Db TrpSerValTyrGluGlyLysValProTyrProArgProGlySerCysAlaSerLysVal 386
1355 TTTGGTGGT---TTTGACTCTCAAAAGGACCTTCTCTGATGATGTTTATAACCTTTGCAAGA 1411
Db AsnGlyGlyArgTyrGlyThrThrLysAspTyrProAspAspAlaIleArgPheAlaArg 406
1412 AGTCATCCAGCCATGATCAATCCAGTGTTCCTTATGAACAATGCCCAATAGTGATCAAA 1471
Db SerHisProLeuMetTyrGlnAlaIleLysProAlaHisLysLysProIleLeuValLys 426
1472 ACGGATGTAATTTACAAATTTACAAATTTGTCGTAGACCGAGTGTGATCGAGAGATGGA 1531
Db ThrAspGlyLysTyrAsnLeuLysGlnIleAlaValAspArgValGluAlaGluAspGly 446
1532 CAGTATGATCTTATGTTTATCGAAACAGATGTTGGACCGCTTCTTAAAGTAGTTCCTCAAT 1591
Db GlnTyrAspValLeuPheIleGlyThrAspAsnGlyIleValLeuLysValIleThrIle 466
1592 CCTAAGGAGACTTGGTATGATTTAGAAGAGTTCCTCTCGAAGAAATGACAGTTTTTCGG 1651
Db TyrAsnGlnGlnMetGluSerMetGluGluValIleLeuGluLeuGlnIlePheLys 486
1652 GAACCGACTGCTATTTACGAATGGAGCTTTTCCATGACGACCAACAATATATATTTGGT 1711
Db AspProValProIleIleSerMetGluIleSerSerLysArgGlnGlnLeuTyrIleGly 506
1712 TCACCGCTGGGGTTCGACGCTCCCTTTTACACCGGTGTGATATTTACGGGAAGCGGTGT 1771
Db SerAlaSerAlaValIleAlaGlnValArgPheHisHisCysAspMetTyrGlySerAlaCys 526
1772 GCTGAGTGTTCCTCCGCCGAGACCTTACTGTGCTTGGAGTGGTTCGTGATGTTCTCGC 1831
Qy

```

Db 527 AlaAspCysLeuAlaArgAspProTyrCysAlaTrpAspGlyLeuSerCysSerArg 546
QY 1832 TATTTTCCCACT-----GCAAGAGACGACCAAGACGACAGATATAAGAAATGGA 1882
Db 547 TyrTyrProThrGlyThrHisAlaLysArgArgPheArgArgGlnAspValArgHisGly 566
QY 1883 GACCCACTGACACACTGTTTCAGACTTACACCATGATTAATACCATGGCCACACCCCTGAA 1942
Db 567 AsnAlaAlaGlnCysPheGlyGlnGlnPheValGlyAspAlaLeuAspLysThrGlu 586
QY 1943 GAGAGATCATCTATGCTAGAGATAGTACACATTTTGGAAATGCGATCGAGTCGAGTCG 2002
Db 587 GluHisLeuAlaTyrGlyIleGluAsnAsnSerThrLeuLeuGluCysThrProArgSer 606
QY 2003 CAGAGCCCTGCTGCTATTGGCAATTCACAGAGCGGAAATGAAGAGCGAAAGAAAGAGATC 2062
Db 607 LeuGlnAlaLysValIleTrpPheValGlnLysGlyArgGluThrArgLysGluGluVal 626
QY 2063 AGATGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGTGTCTACAACAG 2122
Db 627 LysThrAspAspArgValLysMetAspLeuGlyLeuLeuPheLeuArgLeuHisLys 646
QY 2123 AGGATTCAGCAATTAACCTCTGCCATGCGGTGGACATGCGTTCATACAACACTCTTCT 2182
Db 647 SerAspAlaGlyThrTyrPheCysGlnThrValGluHisSerPheValHisThrValArg 666
QY 2183 AAGTAAACCTCGAAGTCAATCAGACAGAGCAATTTGGAAAGACTTCTTCATAAAGATGAT 2242
Db 667 LysIleThrLeuGluValValGluGluLysValGluAspMetPheAsnLysAspAsp 686
QY 2243 GATGAGATGGCTCTAAGACCAA-----GAAATGTCCAATAGCATGACACCTAGCCAG 2296
Db 687 GluGluAspArgHisArgMetProCysProAlaGlnSerSerIleSerGlnGlyAla 706
QY 2297 AGGTCTGTGACAGACTTCATGAGCTCATCAACACCCCACTCAACAGATGGAT 2356
Db 707 LysProTrpTyrLysGluPheLeuGlnLeuIleGlyTyrSerAsnPheGlnArgValGlu 726
QY 2357 GAGTCTGTGAACAAGTTTGGAAAGGACCCAAACACGTCGCAAGGCCAGGACAT 2416
Db 727 GluTyrCysGluLysValTrpCysThrAspArgLysArgLysLysLeuLysMet----- 744
QY 2417 ACCCAGGGAACAGTAACAATGGAAGCACTTACAA-----GAAATGAAGAAAGGTAGAAAC 2473
Db 745 -----SerProSerLysTrpLysTyrAlaAsnProGlnGluLysLysLeuArgSer 761
QY 2474 AGGAGGCCCGCAATTTGAGAGGCCACCCAGG 2506
Db 762 LysProGluHis-----TyrArgLeuProArg 770

```

RESULT 5

```

US-10-408-765A-2003
; Sequence 2003, Application US/10408765A
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Foin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088, 465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2003
; LENGTH: 636
; TYPE: PRT
; ORGANISM: Homo sapiens

```

US-10-408-765A-2003

```

Alignment Scores:
Pred. No.: 3,79e-154 Length: 636
Score: 1917.00 Matches: 359
Percent Similarity: 68.44% Conservative: 103
Best Local Similarity: 53.19% Mismatches: 149
Query Match: 39.15% Indels: 64
DB: 6 Gaps: 9

```

US-09-774-490-1 (1-2709) x US-10-408-765A-2003 (1-636)

```

QY 545 AATTTCATCAAGGTACTTAAGGCATATAATCAGACTCACTTTGTACGCTGTGGAACGGGG 604
Db 2 AsnPheValLysLeuLeuHisAlaTyrAsnArgThrHisLeuLeuAlaCysGlyThrGly 21
QY 605 GCTTTTCATCCAAATTTGCACCTACATTCGAAATGGACATCATCTCGAGGACGACATATTTT 664
Db 22 AlaPheHisProThrCysAlaPheValGluValGlyHisArgAlaGluGluProValLeu 41
QY 665 AAGCTGGAGAACTCACATTTTGAACCGCGTGGAAAGAGTCCATATAGCCCTTAAGCTG 724
Db 42 ArgLeuAspProGlyArgIleGluAspGlyLysSerProTyrAspProArgHis 61
QY 725 CTGACAGCATCCCTTTTAATAGATGGAGAATTATATCTCGAACTGCGAGCTGATTTATG 784
Db 62 ArgAlaAlaSerValLeuValGlyGluLeuTyrSerGlyValAlaAlaAspLeuMet 81
QY 785 GGGGAGACTTTGCTATCTTCGAACTCTTGGGACACACCCCAATCAGGACAGACAG 844
Db 82 GlyArgAspPheThrIlePheArgSerLeuGlyGlnArgProSerLeuArgThrGluPro 101
QY 845 CATGATTCAGGTGGCTCAATGATCCAAAGTTCATTAGTGGCCCACTCATCTCAGAGAT 904
Db 102 HisAspSerArgTrpLeuAsnGluProLysPheValLysValPheTrpIleProGluSer 121
QY 905 GACATCTCTGAAGATGACAAAGTATATCTTTCTTCGTAAGAAATGCAATAGATGAGAA 964
Db 122 GluAsnProAspAspLysIleTyrPhePheArgGluThrAlaValGluAlaAla 141
QY 965 CACTCT---GGAAGACTCAGCTAGATAGATAGATGATGCAAGATGACTTTGGA 1021
Db 142 ProAlaLeuGlyArgLeuSerValSerArgValGlyGlnIleCysArgAsnAspValGly 161
QY 1022 GGGCAGCAAGCTGCTGTAATAAATGGACACATTCCTCAAGCTCGCTGCTGTCACA 1081
Db 162 GlyGlnArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeuValCysSer 181
QY 1082 GTGCCAGGTCCAAATGGCAATTCATTTGATGAATCGCAGGATGATTCCTAATG 1141
Db 182 ValProGlyValGluGly---AspThrHisPheAspGlnLeuGlnAspValPheLeuLeu 200
QY 1142 AACTTTAAGATCCTAAATCCAGTGTATATCGAGTGTGTTTACGATTCAGTAAACATT 1201
Db 201 SerSerArgAspHisArgThrProLeuLeuTyrAlaValPheSerThrSerSerSerile 220
QY 1202 TTCAGGAGTACAGCCGCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1261
Db 221 PheGlnGlySerAlaValCysValTyrSerMetAsnAspValArgArgAlaPheLeuGly 240
QY 1262 CCATATGCCACAGGATGACCCCAACTATCAATGGGTGCGCTTATCAAGGAAGAGTCCCC 1321
Db 241 PropheAlaHisLysGluGlyProMetHisGlnTrpValSerTyrGlnGlyArgValPro 260
QY 1322 TATCCAGGCGCAGAACTTGTCCAGCAAAACATTTGGTGGTGGTGGTGGTGGTGGTGGTGG 1381
Db 261 TyrProArgProGlyMetCysProSerLysThrPheGlyThrPheSerSerThrLysAsp 280
QY 1382 CTTCCTCATGATGATTAACTTTGCAAGAGTCCATCCAGCCCATGATCAATCCAGTGTGT 1441
Db 281 PheProAspAspValIleGlnPheAlaArgAsnHisProLeuMetTyrAsnSerValLeu 300
QY 1442 CCTATGAACAATCGCCCAATAGTATCAAAACCGATGATAAATATCAATTTACACAAATT 1501

```


Db 43 GluTyrPheSerLeuSerHisProLeuLeuAspTyrArgIleLeuLeuMetAspGluAsp 62
QY 395 CGAGTAGCTGTATGTTGGAGCAAGATCATATTTTCATTCGACCTGGTTAATATC 454
Db 63 GlnAspArgIleTyrValGlySerLysAspHisIleLeuSerLeuAsnIleAsnIle 82
QY 455 ---AAGGATTTTCAAAAGATTCTGGCCAGTATCTTACACAGAGAGATCAAGTCAAG 511
Db 83 SerGlnGluAlaLeuSerValPheTrpProAlaSerThrIleLysValGluGluCysLys 102
QY 512 TGGGCTGGAAAAAGACATCTGAAAGAAATGCTTAATTTTCATCAAGCTACTTAAGCATAT 571
Db 103 MetAlaGlyLysAspProThrHisIleGlyCysGlyAsnPheValArgValIleGlnThrPhe 122
QY 572 AATCAGACTCATCTGACCCCTGTGGAGACGGGGCTTTTCATCCAAATTTGCACTACAT 631
Db 123 AsnArgThrHisLeuTyrValCysGlySerGlyAlaPheSerProValCysThrTyrLeu 142
QY 632 GAAATTGGACATCATCTCGAGGACAAATATTTTAAGCTGGAGAACTCACAATTTGAAAC 691
Db 143 AsnArgGlyArgSerGluAspGlnValPheMetIleAsp---SerLysCysGluSer 161
QY 692 GCGCTGGAGAGAGTCCATATGACCTTAAGCTGCTGACAGCATCTCTTTTAATAGATGA 751
Db 162 GlyLysGlyArgCysSerPheAsnProAsnValAsnThrValSerValMetIleAsnGlu 181
QY 752 GAATTATATCTGGAACCTGACCTGATTTTATGGGGCGAGACTTTGCTATCTTCGAACT 811
Db 182 GluLeuPheSerGlyMetTyrIleAspPheMetGlyThrAspAlaAlaIlePheArgSer 201
QY 812 CTGGGGCACCACCCCAATCAGGACAGCAGCATGATTCAGGTGGCTCAATGATCCA 871
Db 202 LeuThrLysArgAsnAlaValArgThrAspGlnHisAsnSerLysTrpLeuSerGluPro 221
QY 872 AAGTTCATTAGTCCCATCTCATCTCAGAGAGTGAATCTCTGAGAGATGACAAAGTATAC 931
Db 222 MetPheValAspAlaHisValIleProAspGlyThrAspProAsnAspAlaLysValTyr 241
QY 932 TTTTCTTCGGTAAATGCAATAGATGAGAGAACACTCTGGAAGAACTACTCAGCTAGA 991
Db 242 PhePhePheLysGluLysLeuThrAspAsnAsnArgSerThrLysGlnIleHisSerMet 261
QY 992 ATAGTCAGATATGCAAGAACTCTTCGGAGGACAGAGTCTGGTGGTGAATAATGGACA 1051
Db 262 IleAlaArgIleCysProAsnAspThrGlyGlyLeuArgSerLeuValAsnLysTrpThr 281
QY 1052 ACATTCTCAAAGCTCGTCTGATTTGCTCAGTCCAGTCCAGTCCAAATGAGCATTCAGACTCAT 1111
Db 282 ThrPheLeuLysAlaArgLeuValCysSerValThrAspGluAspGlyProGluThrHis 301
QY 1112 TTTGATGAAGTGCAGGATGTTCTTCTCATATGCCACAGGATGGAACCTTAAAGATCTCTAAAGATCCCAAGTTGTA 1171
Db 302 PheAspGluLeuGluAspValPheLeuLeuGluThrAspAsnProArgThrLeuVal 321
QY 1172 TATGAGTGTTCAGACTTCCAGTAAACATTTTCAAGGATCAGCGCTGTGTATGATAGC 1231
Db 322 TyrGlyIlePheThrSerSerSerValPheLysGlySerAlaValCysValTyrHis 341
QY 1232 ATGAGTGTATGTAGAGAGGTGTTCTCTGCTCATATGCCACAGGATGGAACCTCAAT 1291
Db 342 LeuSerAspIleGlnThrValPheAsnGlyProPheAlaHisLysGluGlyProAsnHis 361
QY 1292 CAATGGGTGCTTATCAAGGAGAGTCCCTCATCCAGCGCCAGGAACCTGTGCCAGCAAA 1351
Db 362 GlnLeuIleSerTyrGlnGlyArgIleProTyrProArgProGlyThrCysProGlyGly 381
QY 1352 ACATTT---GGTGGTTTGCATCTCAAGAGACTTCTGTATGATGATGTTAATCACTTGA 1408
Db 382 AlaPheThrProAsnMetArgThrThrLysGluPheProAspAspValValThrPheIle 401
QY 1409 AGAAGTCATCCAGCCATGTCATCAACCTGTTCTCATGAACTCCGCCAATAGTATGC 1468
Db 402 ArgAsnHisProLeuMetTyrAsnSerIleTyrProIleHisLysArgProLeuIleVal 421

QY 1469 AAAACGGAGTAAATTTATCAATTTTACACAAATTTCTGTAGACCGAGTGGATGCAGAGAT 1528
Db 422 ArgIleGlyThrAspTyrLysTyrThrLysIleAlaValAspArgValAsnAlaAlaAsp 441
QY 1529 GGACAGTATGATGATGTTATCGGAACAGATGTTGGACCGTCTTAAAGTAGTTTCA 1588
Db 442 GlyArgTyrHisValLeuPheLeuGlyThrAspArgGlyThrValGlnLysValValVal 461
QY 1589 ATTCTAAGGAGACTTGGTATGATTAGACAGGTTCTGCTGGAAGAAATCAGACTTTT 1648
Db 462 LeuProThrAsnAsnSerValSer---GlyGluLeuIleLeuGluLeuGluValPhe 480
QY 1649 CGGAACCGACTGCTATTTTCAGCAATGAGATGCTTCCATTAAGCAGCAACAATATATT 1708
Db 481 LysAsnHisAlaProIleThrThrMetLysIleSerSerLysLysGluGlu----- 497
QY 1709 GGTTCACCGCTGGGTTGCCAGCTCCCTTTTACACCGGTGTGATATTAT 1759
Db 498 -----ProLysThrArgCysGluThrTrp 505
RESULT 7
US-10-369-072-97
; Sequence 97, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkes, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine B
; APPLICANT: Shinkets, Richard
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Domain search
; OTHER INFORMATION: result
US-10-369-072-97

Alignment Scores:
Pred. No.:      2,65e-88      Length:      431
Score:          1136.50      Matches:      229
Percent Similarity: 66.52%      Conservative: 69
Best Local Similarity: 51.12%      Mismatches: 125
Query Match:      23.21%      Indels:      25
DB:              6          Gaps:      9

US-09-774-490-1 (1-2709) x US-10-369-072-97 (1-431)
QY 368 TATCATACCTCTCTTTGGATGAGAAACGAGTAGCTGTATGTTGGAGCAAGATCAC 427
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 PheArgThrLeuLeuAspGluAspArgGlyArgLeuTyrValGlyAlaArgAsnHis 20
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 428 ATATTTCATTCGACCTGGTTAATATCATCAAGGATTTTCAA-----AAGATT 472
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21 ValTyr-----ValLeuAsnLeuGluAspLeuSerGluValLeuAsnLeuLysIle 37
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 473 GTGTGCCAGATATCTTACACCAAGAGATGAATGCAAGTGGGCTCGGAAAAGACATCTG 532
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 38 GlyTyrProAlaSerCysGluLysCysGluGluCysAsnMetLysGlyLysSerProLeu 57
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 533 AAAGATGTGCTAATTCATCAAGGTACTTAAGGATATATATCAGACATCACTGTACGCC 592
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 58 ThrGluCysThrAsnPheIleArgValLeuGlnAlaTyrAsnAspThrHisLeuTyrVal 77
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 593 TGTGGAACGGGGCTTTTCATCCAAATTTGCACCTACATTGAATTTGGACATCATCTGAG 652
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 78 CysGlyThrAsnAlaPheAsnProLysCysThrLeuIleAsnLeuGly----- 93
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 653 GACAATATTTTAAAGTCGAGAACATCACATTTTGAACCGCGCTGGGAAGTAGTCCATAT 712
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 94 ---AspLeuPheSerLeuAspAsnAspAsnGluGluSerGlyCysGlyAspCysProTyr 112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 713 GACCTTAAGCTGCTGACGACATCCCTTTTAAATAGATGGAGAATATATCTTGAAGACTGCA 772
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 AspProLeuGlyAsnThrThrSerValLeuValGlyGlyGluLeuTyrSerGlyThrAla 132
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 773 GCTGATTTTATCGGCGGACATTTGCTATCTTCCGAATCTTGGGCACACACCCCAATC 832
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 133 IleAspPheSerGlyArgAspProSerIleArgGLeuLeuGlySerHisAspGlyLeu 152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 833 AGGACAGACGACATGATTCACAGTGGCTCAATGATCCAAAGTTCATTAGTGCACCCACCTC 892
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 153 ArgThrGluPheHisAspSerLysTrpLeuAsnLeuProAsnPheValAspSerTyrPro 172
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 893 ATCTCAGAGATGACATCTCCTGAAGATGACAAAGTATATCTTTCTTCGTGGAATGCA 952
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 173 IleHisTyrValHisSerPheSerAspLysValTyrPhePheArgGluThrAla 192
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 953 ATAGATGGAGACATCTCTGCA---AAAGCTACTCAGCTAGATAGGTGAGATGCAAG 1009
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 193 Val-----GluTyrSerAsnCysLysAlaIleHisSerArgValAlaArgValCysLys 210
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1010 AATGACTTTTGGGGGCACAGAAGTCTGGTGAATAAATGGAACAACATTTCTCAAAGCTCGT 1069
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 211 AsnAspProGlyGlyHisSerTyrLeuGluAsnLysTrpThrPheLeuLysAlaArg 230
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1070 CTGATTTGCTAGTGCAGGTCCAAATGGCATTGACATCATTTTGTGAACTGACGAGAT 1129
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 231 LeuAsnCysSerIleProGly---GluGlyThrProPheTyrPheAsnGluLeuGlnAla 249
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1130 GTATTCTTAATCACTTTAAAGATCCTAAATCCAGTTGTATATGGAGTGTTCAGACT 1189
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 250 AlaPheLeuProThrAspAsnAspThrAspProValLeuTyrGlyValPheThrThr 269
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1190 TCCAGTAAATTTTCAAGGGATCAGCCGTGTGTATGTATAGCATGAGTGTGAGGAGG 1249
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 8

```

US-10-369-072-47
; Sequence 47, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zehusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkes, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10369,072
; PRIORITY FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07

```

```

Db 270 SerSerAsnSerSerAlaGlySerAlaValCysAlaPheSerMetLysAspIleAsnGln 289
QY 1250 GTGTTCTCTTGGTCCATATGCCACAGGATGAGCCCACTATCAATCAATGGTCCCTTATCAA 1309
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 290 ValPheGluGlyProPheLysHisGln---GlyProAsnSerLysTrpLeuProTyrArg 308
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1310 GGAAGAGTCCCTATCCACGCGCCAGGAACCTTGTCACCAACAAACATTTGGTGGTTTTCAC 1369
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 309 GlyArgValProTyrProArgProGlyGlnCysProAsnSerSerAsnGly----- 325
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1370 TCTACAAAGGACCTCTCTGATGATGATTATACCTTTTGCACAAAGTCATCCAGCCCATGTAC 1429
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 -----AspLeuProAspAspThrLeuAsnPheIleArgCysHisProLeuMetAsp 342
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1430 ANTCCAGTGTTCATGAACAATCGCCCAATAGTATGATCAAAACGGATGTAATATATCAA 1489
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 343 AspValValProProLeuHisAsnValProLeuPheValGlyGlnSerGlyAsnTyrArg 362
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1490 TTTACACAAATTTGCTAGACCGGATGCGATGCGACAGATGCGAAGATGGACAG---TATGATGTTATG 1546
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 363 LeuThrSerIleAlaValAspArgValGluAlaGlyAspGlyGlnIleTyrThrValLeu 382
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1547 TTTATCGCAACAGATGTTGGACCGCTTCTTAAAGTAGTTTCAATCTCTAAGGACACTGG 1606
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 383 PheLeuGlyThrAspAspGlyArgValLeuLysGlnValValLeuSerArgSerSer 402
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1607 TATGATTAGAAGAGTCTCTCTGGAAGAAATGACAGTTTTTCGGGACCGACTGCTATT 1666
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 403 AlaSerTyrGluValValValLeuGluGluSerLeuValPheProAspGlyGluProIle 422
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1667 TCAGCAATGGAGCTTTCACCTAAG 1690
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 423 GlnAsnMetGluIleSerGlnLys 430
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

; Remaining Prior Application data removed - See File Wrapper or PALM.

US-09-774-490-1 (1-2709) x US-10-369-072-47 (1-834)

Db 212 GluProHisPheValGlySerAlaPheValProGluSerValGlySerPheThrGlyAsp 231


```
Db 571 GlyThrAspLeuValLeuProCysHisLeuSerAsnLeuAlaHisIstPrThr 590
QY 2027 TTC-----CAGAGCGAAATGAAGCGAAGAGAGATCAGAGTGGATGATCATATC 2080
Db 591 PheGlySerGlnAspLeuProAlaGluGlnProGlySerPheLeuTyraAspThrGlyLeu 610
QY 2081 ATCAGGACAGATCAGAGCGCTTCTGCTACGTAGTCTACACAGAGGATTCAGGCAATTAC 2140
Db 611 -----GlnAlaLeuValMetAlaAlaGlnSerArgHisSerGlyProTyr 626
QY 2141 CTCCTGCATCGCGTGGAAATCGG 2164
Db 627 ArgCysTyrSerGluGluGlnGly 634
```

RESULT 9

```
US-10-369-072-91
; Sequence 91, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsebrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Fatturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shinkets, Richard
; APPLICANT: Padigar, Muralidhara
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; PRIOR FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 91
; LENGTH: 430
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Domain search
; OTHER INFORMATION: result
US-10-369-072-91
Alignment Scores:
```

```
Pred. No.: 9.67e-74 Length: 430
Score: 964.00 Matches: 204
Percent Similarity: 62.42% Conservative: 75
Best Local Similarity: 45.84% Mismatches: 138
Query Match: 19.69% Indels: 30
DB: 6 Gaps: 10

US-09-774-490-1 (1-2709) * US-10-369-072-91 (1-430)
QY 380 CTTTGTGATGAGGAACGAGTAGGCTGTATGTTGGAGCAAGGATCATATTTTCATTC 439
Db 5 LeuLeuAspGluAspAsnGlyThrLeuTyrValGlyAlaArgAsnArgLeuTyrAlaLeu 24
QY 440 GACCTGGTTAATATCAAGGATTTTCAAAGATTGTGTGGCCAGATCTTTACACCAGAAGA 499
Db 25 SerLeuAsnLeuIleSerGluAlaGluValLysThrGlyProValSerSerProAsp 44
QY 500 ---GATGAATGCAAGTGGGCTGGAAAGACATCTCTGAAAGAAATGTGTAATTTTCATCAAG 556
Db 45 CysGluGluCysValSerLysGlyLysAspProProThrAspCysGlnAsnPheIleArg 64
QY 557 GTACTTAAGGCATATAATCAGACTCATTGTACGCTGTGTGGAACGGGGGCTTTTCATCCA 616
Db 65 LeuLeuLeuAspTyrAsnAlaAspArgLeuLeuValCysGlyThrAsnAlaPheGlnPro 84
QY 617 ATTTGCACCTACATTGAAATTGGACATCATCTCAGGACATATTTTAAAGCTGGGAAC 676
Db 85 ValCysArgLeuIleAsnLeuGly-----AsnLeuAspArgLeuGluVal 99
QY 677 TCACATTTTGAAGACGGCGGTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCC 736
Db 100 GlyArg---GluSerGlyArgGlyArgCysProTyrAspProGlnHisAsnSerThrAla 118
QY 737 CTTTAAATAGATGAGAAATATATCTCTGGAACCTGAGCTGATTTTATGCGGGGAGACTTT 796
Db 119 ValLeuValAspGlyGluLeuTyrValGlyThrValAlaAspPheSerGlySerAspPro 138
QY 797 GCTATCTTCGAACTCTT-----GGGCACCACCACCA---ATCAGG 835
Db 139 AlaIleTyrArgSerLeuSerValArgArgLeuLysGlyThrSerGlyProSerLeuArg 158
QY 836 ACAGAGCAGCATGATTCAGGTGGCTCAATGATCAAAAGTTTCATTAGTGGCCACCTCATC 895
Db 159 ThrValLeuTyrAspSerArgTyrLeuAsnGluProAsnPheValTyrAlaPhe----- 176
QY 896 TCAGAGAGTGACAATCTGAAGATGACAAAGTATATTTTCTCCGTGGAATAATGCAATA 955
Db 177 -----GluSerGlyAspPheValTyrPhePheArgGluThrAlaVal 191
QY 956 GATGAGAGCACTCTGGAAAGCTACTCAGCTAGAATAGTTCAGATATCAAGAAATCAC 1015
Db 192 GluAspGluAsnGlyLysAlaTyrValSerArgValAlaArgValCysLysAsnAsp 211
QY 1016 TTTGGAGGCGACAGAAGTCTGGTGAATAAATGGAACAACATTCCTCAAGCTCGTCTGATT 1075
Db 212 ValGlyGlyProArgSerLeuSerLysIleThrPhePheLeuLysAlaArgLeuGlu 231
QY 1076 TGCTCAGTGGCAGTCCAAATGGCATTGACACTCTTTTGTATGAACTGCAGAGTGTATTC 1135
Db 232 CysSerValProGlyGluSerProPhe-----TyrPheAsnGluLeuGlnAlaAlaPhe 249
QY 1136 CTAATGAACCTTAAAGATCCTAAAAATCCAGTCTGTATATGAGTGTGTATACCACTTCCAGT 1195
Db 250 LeuLeuProAlaGlySerGluSerAspValLeuTyrGlyValPheSerThrSerSer 269
QY 1196 AACATTTTCAAGGATCAGCGCTGTGTATGTATAGCATGTAGTGTAGAGAGGAGGTGTC 1255
Db 270 AsnSerIleProGlySerAlaValCysAlaPheSerLeuSerAspIleAsnAlaValPhe 289
QY 1256 CTTGGTCCATATGCCACAGGATGGACCCCAACTATCAATGGTGGCTTAT---CAAGGA 1312
Db 290 AsnGluProPheLysGluCysGluThrGlyAsnSerGlnTyrLeuProTyrProArgGly 309
```



```

QY 1313 AGAGTCCCTATCCAGCGCCAGAACTTGTCCAGCAAAACATTTGGTGGTTTGTACTCT 1372
Db 310 LeuValProPheProArgProGlyThrCysProAsnAsnSerLeu-----Ser 325
QY 1373 ACAAGGACCTTCTGATGATGTATACTTTTGCAGAGATCATCCAGCCATGTACAT 1432
Db 326 SerLysAspLeuProAspThrLeuAsnPheIleLysThrHisProLeuMetAspGlu 345
QY 1433 CCAGTGTTCCTATGACCAATCGCCCAATAGTATGATCAAAACCGATGATAATTATCAATTT 1492
Db 346 AlaValGlnProLeuThrGlyArgProLeuLeuValLysThrAspSerAsnTyrLeuLeu 365
QY 1493 ACACAAATTTGCTAGACCGAGTGGATGACAGAGATGACAGATGATGATGATGATGATC 1552
Db 366 ThrSerIleAlaValAspArgValArgThrAspGlyGlyAsnTyrThrValLeuPheLeu 385
QY 1553 GGAACAGATGTTGGGACCGTCTTAAAGTAGTTTCAATTCCTTAAGGAGAGATGATGAT 1612
Db 386 GlyThrSerAspGlyArgIleLeuLysValValLeu-----SerGluSerSerSer 403
QY 1613 TTAGAAGAGTTCTGCTGGAAGAAATGACAGATTTTTCGGGAACCCGACTGCTATTTTCAGCA 1672
Db 404 SerGluSerValValLeuGluGluIleSerValPheProGlySerProIleSerAsp 423
QY 1673 ATGGAGCTTTCCTACTAAGCAG 1693
Db 424 LeuValIleSerProLysLys 430

RESULT 10
PCT-US03-09929-10
; Sequence 10, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 10
; LENGTH: 638
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-09929-10

Alignment Scores:
Pred. No.: 1.02e-72 Length: 638
Score: 952.50 Matches: 230
Percent Similarity: 50.92% Conservative: 103
Best Local Similarity: 35.17% Mismatches: 250
Query Match: 19.45% Indels: 71
DB: 1 Gaps: 19

```

```

US-09-774-490-1 (1-2709) x PCT-US03-09929-10 (1-638)
QY 278 AACATGTGCAAGGCTGAAATTTATCTTACAAAGAAATGTTTGAATCCAAACATGTGATC 337
Db 8 AsnLeuValProArgLysThrValSerSerGlyGluLeuAlaThrValValArgPhe 27
QY 338 ACTTTCATATGCTTGGCCCAACAGCTCCAGTTATCATATCATCTTCTTTTGGATGAGGAACGG 397
Db 28 SerGlnThrGlyIle-----GlnAspPheLeuThrLeuThrLeuGluProThr 44
QY 398 AGTAGGCTGTATGTTGGAGCAAGGATCACATATTTTCATTCGACCTGGTTAATATCAAG 457
Db 45 GlyLeuLeuTyrValGlyAlaArgGluAlaLeuPheAlaPheSerMetGluAlaLeuGlu 64
QY 458 GATTTTCAAAAGATGTGTGCGCCAGTATCTTACACCAAGAGATGAATGCAAGTGGGCT 517
Db 65 LeuGlnGlyAlaIleSerTyrPheValGluProValGluLysThrGluCysIleGlnLys 84
QY 518 GGAAGAAGATCTCTGAAAGAAATGTCTAATTTTCATCAAGGTACTTTAAGGCATATATATCAG 577
Db 85 GlyLysAsnSerGlnThrGluCysPheAsnPheIleArgPheLeuGlnProTyrAsnAla 104
QY 578 ACTCACTTGTAGCTGTGGAACCGGGGCTTTTCATCCAAATTTGACCTACATTTGAAAT 637
Db 105 SerHisLeuTyrValCysGlyThrTyrAlaPheGlnProLysCysThrTyrValAsnMet 124
QY 638 GGACATCATCTCGAGGACAATATTTTAAAGCTGGAGNACTCACATTTTGAACACGGCGT 697
Db 125 -----LeuThrPheThrLeuGluHisGlyGluPheGluAspGlyLys 138
QY 698 GGAAGAGTCCATATGACCTTAAGCTGTGACAGCATCCCTTTTAAATAGATGGAGAATTA 757
Db 139 GlyLysCysProTyrAspProAlaLysGlyHisAlaGlyLeuLeuValAspGlyGluLeu 158
QY 758 TACTCTGGAATGACGCTGATTTTATGGGCGAGACTTTGCTATCTTCGGAATCTTGGG 817
Db 159 TyrSerAlaThrLeuAsnAsnPheLeuGlyThrGluProIleIleLeuArgAsnMetGly 178
QY 818 CACCACCAACCAATCAGGACAGAGCAGCATGATTCAGCTGGCTCAATGATCCAAAGTTC 877
Db 179 ProHisHisSerMetLysThrGlu---TyrLeuAlaPheThrPheLysGluProHisPhe 197
QY 878 ATTAGTGCCACCTCATCTCAGAGAGTGCACAATCTCT-----GAAGATGACAAAGTA 928
Db 198 ValGlySerAlaTyrValProGluSerValGlySerPheThrGlyAspAspLysVal 217
QY 929 TACTTTTCTTCGTTGAAATGCAATAGATGGAGAACACTCTGGAAAAGCTACTCACGCT 988
Db 218 TyrPhePheArgGluArgAlaValGluSerAspCysTyrAlaGluGlnValAla 237
QY 989 AGAATAGTCAAGATATCAAGAATGACTTTTGGAGGACACAGTCTGGTGAATAAATGG 1048
Db 238 ArgValAlaArgValCysLysGlyAspMetGlyGlyAlaArgThrLeuGlnArgLysTrp 257
QY 1049 ACAACATTTCTCAAAAGCTCGTCTGATTTTTCAGTGCAGCTCCAGGTCCAAATGCGACACT 1108
Db 258 ThrThrPheLeuLysAlaArgLeuAlaCysSerAlaPro-----AsnTyrGlnLeu 274
QY 1109 CATTTTGATGAACCTGAGGATGATTCCTAATGAACCTTTAAAGATCTTAAATATCCAGTT 1168
Db 275 TyrPheAsnGlnLeuGlnAlaMetHisThrLeuGlnAspThrSerTyrPheAsnThrThr 294
QY 1169 GTATATGAGTGTTCAGGACTTCCAGTAAACATTTTCAAGGATCAGCGCTGTATGTAT 1228
Db 295 PheGlyValPheGlnAlaGlnTrpGlyAspMetTyrLeuSerAlaIleCysGluTyr 314
QY 1229 AGCATGATGATGTGAGAAGGGTGTCTCTTGGTCCATATGCCAGGGATGGACCCCAAC 1288
Db 315 GlnLeuGluGluIleGlnArgValPheGluGlyProTyrLysGluTyrHisGluGluAla 334
QY 1289 TATCAATGGGTGCTTATCAAGGAAGAGTCCCTTATCCAGGCCAGGAACCTTGTCCACG 1348
Db 1289 TATCAATGGGTGCTTATCAAGGAAGAGTCCCTTATCCAGGCCAGGAACCTTGTCCACG 1348

```

```

Db 335 GlnLysTrpAspArgTyrThrAspProValProSerProArgProGlySerCysIleAsn 354
Qy 1349 AAAACATTT-----GGTGGTTTGTACTCTCAAGAGGACCTTCCTGATCATGCTTTATATACC 1402
Db 355 AsnTrpHisArgHisGlyTyrThrSerSerLeuGluLeuProAsnIleLeuAsn 374
Qy 1403 TTTGCAAGAGTCAATCCAGCCATGTACATCCAGTGTTCCTTATGAACAATCGCCCAATA 1462
Db 375 PheValLysLysHisProLeuMetGluGluGlnValGlyProArgTrpSerArgProLeu 394
Qy 1463 GTGATCAAAACGGATGTAATATTCAATTTACAAATTTGCTAGACCGAGTGGATGCA 1522
Db 395 LeuValLysLysGlyThrAsn-----PheThrHisLeuValAlaAspArgValThrGly 412
Qy 1523 GAAGATGGA---CAGTATCATGTTATGTTTATCGGAACAGATGTTGGACCCCTTCTTAA 1579
Db 413 LeuAspGlyAlaThrTyrThrValLeuPheIleGlyThrGlyAspGlyTrpLeuLeuLys 432
Qy 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAAGAGTTCCTGCTGGAAGAAATG 1639
Db 433 AlaValSerLeu-----GlyProTrpValHis-----LeuIleGluGluLeu 446
Qy 1640 ACAGTTTTT---CGGAAACCGACTGCTATTTACGAATAGAGCTTTCCACTAAGCAGCAA 1696
Db 447 GlnLeuPheAspGlnPro-----MetArgSerLeuValLeuSerGlnSerLysLys 464
Qy 1697 CAATATATATTTGGTTCAACGGCTGGGTTGCCAGCTCCCTTTACACCGGTGCTATATT 1756
Db 465 LeuLeuPheAlaGlySerArgSerGlnLeuValGlnLeuProValAlaAlaAspCysMetLys 484
Qy 1757 TACGGAAACCGTGTCTCAGTGTTCCTCGCCGAGACCCCTTACTGTCTGGTGGATGCT 1816
Db 485 Tyr---ArgSerCysAlaAspCysValLeuAlaArgAspProTyrCysAlaTrpSerVal 503
Qy 1817 TCTGCATGTTCTCGCTATTTTCCCACTGCAAGAGACGCAAGACGA-----1864
Db 504 Asn-----ThrSerArgCysValAlaVal 511
Qy 1865 -----CAAGATATAAGAAATGGAGACCCACTGACTCAC 1897
Db 512 GlyGlyHisSerGlySerLeuLeuIleGlnHisValMetThrSerAspThrSerGlyIle 531
Qy 1898 TGTTCAGACTTACACCATGATTAATCACCATGCCACAGC-----CCTCAAGAGAGA 1948
Db 532 Cys-----AsnLeuArgGlySerLysValArgProThrProLys 545
Qy 1949 ATCATCTATGTTGATAGAAATAGTAGCACATTTTGGAAATGAGTCCGAGTCCGAGAGA 2008
Db 546 AsnIleThrValValAlaGlyThrAspLeuValLeuProCysHisLeuSerSerAsnLeu 565
Qy 2009 GCCTGCTCTATTGGCAATTCCAGAGCGCAAT-----GAAGAGCGAAAGAGAGATC 2062
Db 566 AlaHisAlaArgTrpThrPheGlyGlyArgAspLeuProAlaGluGlnProGlySerPhe 585
Qy 2063 AGAGTGGATCATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAGTCTACACAG 2122
Db 586 LeuTyrAspAlaArgLeu-----GlnAlaLeuValValMetAlaAlaGlnPro 601
Qy 2123 AAGATTACGCAATTTACTCTGCCATGCGGTGGAAACATGGG 2164
Db 602 ArgHisAlaGlyAlaTyrHisCysPheSerGluGluGlnGly 615

RESULT 11
US-10-369-072-14
; Sequence 14, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
;
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigaru, Muralidhara
;
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 14
; LENGTH: 833
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-072-14

Alignment Scores:
Pred. No.: 1,11e-72 Length: 833
Score: 952.50 Matches: 236
Percent Similarity: 50.82% Conservative: 106
Best Local Similarity: 35.07% Mismatches: 264
Query Match: 19.45% Indels: 67
DB: 6 Gaps: 22

US-09-774-490-1 (1-2709) x US-10-369-072-14 (1-833)
Qy 206 TGGTTAACTAGATGTGTCTTTCTGGGGAGTATTACTTACGACGAGCAACTAT 265
Db 8 TrpLeu-----LeuAlaAlaArgLeuTrpGlyLeuGlyAlaGluValTrpTrp 25
Qy 266 CAGATCGGAAGACAAATGTCCAGGCTGAAATTCCTCAAGAAATGTTGGAATCC 325
Db 26 -----AsnLeuValProArgLysThrValSerSerGlyGluLeuAlaThrVal 41
Qy 326 AACAAATGTGATCACTTTTCAATGGCTTGGCCAAACAGCTCCAGTTCATPACCTTCTTTTG 385
Db 42 ValArgArgPheSerGlnThrGlyIle-----GlnAspPheLeuThrLeuThrLeu 58
Qy 386 GATGAGAACCGAGTAGGCTGTATGTTGGACGAAGAGTACATATTTCATTTCGACCTG 445
Db 59 ThrGluProThrGlyLeuLeuTyrValGlyAlaArgGluAlaLeuPheAlaPheSerVal 78
Qy 446 GTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACACCAAGAGAGATCAA 505
Db 79 GluAlaLeuGluLeuGlnGlyAlaIleSerTrpGluAlaProValGluLysLysThrGlu 98
```

506 TCGAAGTGGCTGGAAAGACATCTCTGAAGAAATGGCTAAATTCATCAAGTACTTAAG 565
Db CysIleGlnLysGlyAsnGlnThrGluCysPheAsnPheIleArgPheLeuGln 118
566 GCATATAATCAGACTCAGCTGACGCTGTGGAACGGGGCTTTTCATCAATTTGCACC 625
Db ProTyrAsnAlaSerHisIleuValCysGlyThrTyrAlaPheGlnProLysCysThr 138
626 TCATTTGAAATGGACATCATCTCGAGCAATATATTTTAAAGCTGGAGACATTCATTT 685
Db TyrValAsnMet-----LeuThrPheThrLeuGluHisGlyGluPhe 152
686 GAAACGGCTGGGAGAGTCCATATGACCCCTAGCTCTGACAGCATCCCTTTTAATA 745
Db GluAspGlyLeuGlyCysProTyrAspProAlaLysGlyHisAlaGlyLeuLeuVal 172
746 GATGAGAAATATATCTCTGCACTGCGAGCTGATTTTATGGGCGGAGACTTCTCATCTTC 805
Db AspGlyGluLeuTyrSerAlaThrLeuAsnAsnPheLeuGlyThrGluProIleLeuLeu 192
806 CGAATCTTTGGCACCCACCAATCAGGACAGCAGCATGATTCAGGCTGGCTCAAT 865
Db ArgAsnMetGlyProHisHisSerMetLysThrGlu---TyrLeuAlaPheThrLeuAsn 211
866 GATCMAAGTTCATGATGCCACCTCATCTCAGAGAGTGACAACTCT-----GAA 916
Db GluProHisPheValGlySerAlaTyrValProGluSerValGlySerPheThrGlyAsp 231
917 GATGACAAAGTATCTTTCTTCGCGTGAATGCAATAGATGAGGAGACACTCTGGAAA 976
Db AspAspLysValTyrPheLeuPheArgGluArgAlaValGluSerAspCysTyrAlaGlu 251
977 GCTACTCAGCTAGATAGTTCAGATATCAGAAATGACTTTGGAGGGCAGCAAGTCTG 1036
Db GlnValValAlaArgValAlaArgValCysLysGlyAspMetGlyAlaArgThrLeu 271
1037 GTGAATAATGACAACTCTCCTCAAGCTGCTGATGTTGCTGATGCTGAGTCCAGGTC 1096
Db GlnArgLysThrThrPheLeuLysAlaArgLeuAlaCysSerAlaPro----- 288
1097 GGCATTGACACTCATTTGATCACTGACGAGTATCTTCTTAATGAACCTTTAAGATCCT 1156
Db AsnThrGlnLeuTyrPheAsnGlnLeuGlnAlaMetHisThrLeuGlnAspThrSerTrp 308
1157 AAAAAATCCAGTTGTATGAGAGTGTTCAGACTTCCAGTAAACATTTTCAAGGATCAGCC 1216
Db HisAsnThrThrPhePheGlyValPheGlnAlaGlnTrpGlyAspMetTyrLeuSerAla 328
1217 GTGTGTATGATAGATGATGATGAGAGGGTGTCTTGTGCTCATATGCCCCACAGG 1276
Db IleCysGluTyrGlnLeuGluGluIleGlnArgValPheGluGlyProTyrLysGluTyr 348
1277 GATGACCCCAATCATCAATGGTGGCTTATCAAGNAGAGTCCCTATCCAGGCGCAGA 1336
Db HisGluGluAlaGlnLysTrpAspArgfyrThrAspProValProSerProArgProGly 368
1337 ACTTGTCCCAAGAAAACATTT-----GGTGGTTTGTACTACAAAGACCTTCTCTGAT 1390
Db SerCysIleAsnAsnTrpHisArgHisGlyThrThrSerSerLeuGluLeuProAsp 388
1391 GATGTTATACCTTTTGCAGAGAGTCAATCCAGCATGTAATCCAGTGTTCCTATGAAAC 1450
Db AsnIleLeuAsnPheValLysLysHisProLeuMetGluGlnValGlyProArgTrp 408
1451 AATCCCAATAGTATCAAAACGGATGTAATATCAATTTTACAAATTTGCTGATAC 1510
Db SerArgProLeuLeuValLysLysGlyThrAsn-----PheThrHisLeuValAlaAsp 426
1511 CGAGTGGATGCAGAGACGGA---CAGTATGATGTTATGTTTATCGGAACAGATCTGGG 1567
Db ArgValThrGlyLeuAspGlyAlaThrTyrThrValLeuPheIleGlyThrGlyGlnAla 446
1568 ACCGTTCTTAAAGTAGTTTCAATCTCCTAAGGAGACTTGTGATGATTTAGAGAGGTTCTG 1627

447 TrpLeuLeuLysAlaValSerLeu-----GlyProTrpValHis-----Leu 460
1628 CTGGAAGAAATGACAGTGTTTT---CGGAAACCGAGTGTATTTTCAGCAATGAGCTTCC 1684
Db IleGluGluLeuGlnLeuPheAspGlnGluPro-----MetArgSerLeuValLeuSer 478
1685 ACTAAGCAGCAACACTATATATTTGGTTCAACGGCTGGGTTGCCAGCTCCCTTTTACAC 1744
Db GlnSerGlnLysLeuLeuPheAlaGlySerArgSerGlnLeuValGlnLeuProValAla 498
1745 CGGTGTGATATTTACGGGAAACCGTGTCTGAGTGTTCCTCGCCCGGAGACCTTACTGT 1804
Db AspCysMetLysTyr---ArgSerCysAlaAspCysValLeuAlaArgAspProTyrCys 517
1805 GCTTGG-----GATGGTTCTGCATGT-----TCTCGCTATTTTCCC 1840
Db AlaTrpSerValAsnThrSerArgCysValAlaValGlyGlyHisSerGlySerPheLeu 537
1841 ACTCAAGACAGCAGCAAGACAGATATAGAAATGAGAGACCCACTGACTCACTGT 1900
Db Ile-----GlnHisValMetThrSerAspThrSerGlyIleCys 550
1901 TCAGACTTACCATGATATACCATCGCCACAGC-----CCTGAAGAGAGATC 1951
Db AsnLeuArgGlySerLysLysValArgProThrProLysAsn 564
1952 ATCTATGTTAGAGATAGTAGCACATTTTGGATGTCAGTCCGAAAGTCCGAGAGAGG 2011
Db IleThrValValAlaGlyThrAspLeuValLeuProCysHisLeuSerSerAsnLeuAla 584
2012 CTGTCTTATTCGCAATTCAGAGGCGAAAT-----GAAGAGCGAAAGAGAGATCAGA 2065
Db HisAlaArgTrpThrPheGlyGlyArgAspLeuProAlaGluGlnProGlySerPheLeu 604
2066 GTGATGATCATATCATCATCGGACAGATCAAGGCTTCTGCTAGCTAGTCTTCAACAGAG 2125
Db TyrAspAlaArgLeu-----GlnAlaLeuValMetAlaAlaGlnProArg 620
2126 GATTCAAGCAATTAACCTCTGTCATCGCGTGAACATGGG 2164
Db HisAlaGlyAlaTyrHisCysPheSerGluGlnGly 633

RESULT 12
PCT-US03-09929-6
; Sequence 6, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 179
 ; SOFTWARE: CuraseqList version 0.1
 ; SEQ ID NO 6
 ; LENGTH: 833
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 PCT-US03-09929-6

Alignment Scores:
 Pred. No.: 1.63e-72 Length: 833
 Score: 950.50 Matches: 235
 Percent Similarity: 50.06% Conservative: 104
 Best Local Similarity: 34.66% Mismatches: 262
 Query Match: 19.41% Indels: 77
 DB: 1 Gaps: 21

US-09-774-490-1 (1-2709) x PCT-US03-09929-6 (1-833)

QY 206 TGGTAACTAGGATTGCTGCTCTTTCTGGGAGTATTACTTACAGCAAGAGCAAACTAT 265
 DB 8 TrpLeu-----LeuAlaAlaArgLeuTrpGlyLeuGlyLeuGlyLeuValTrpTrp 25
 QY 266 CAGAAATGGAAGAACAAATGTGCCAGGCTGAAATATCTCACAAGAAATGTGGAATCC 325
 DB 26 -----AsnLeuValProArgLysThrValSerSerGlyGluLeuAlaThrVal 41
 QY 326 AACATGTGATCATCTTCAATGGCTTGGCCACAGCTCCAGTTATCATCTCTCTTTTG 385
 DB 42 ValArgArgPheSerGlnThrGlyLe-----GlnAspPheLeuThrLeuThrLeu 58
 QY 386 GATGAGGAACGAGTAGGCTGTATGTTGGAGCAAGAGTACATATTTTCATTCGACCTG 445
 DB 59 ThrGluProThrGlyLeuLeuLeuValGlyAlaAlaArgAspHisAlaSerAlaLeuGlyVal 78
 QY 446 GTTAATATCAAGGATTTTCAAAGAGTGTGTGCCAGTATCTTACACCAAGAGAGATGAA 505
 DB 79 ProValLeuLeuGlnAlaValIleSerTrpGluAlaProValGluLysLysThrGlu 98
 QY 506 TCGAAGTGGCTGGAAAGACATCTCGAAAGATGTGCTAAATTTTCATCAAGTACTTAAG 565
 DB 99 CysIleGlnLysGlyLysAsnGlnThrGluCysPheAsnPheIleArgPheLeuGln 118
 QY 566 GCATATAATCAGATCAGCTGTGAGCTGTGGAACGGGGCTTTTTCATCCAATTTGACCC 625
 DB 119 ProTyrAsnAlaSerHisLeuTrpValCysGlyThrThrAlaPheGlnProLysCysThr 138
 QY 626 TCATTTGAAATGGACATCATCTGTAGGACAATATTTTAACTGGAGAACTCAATTTT 685
 DB 139 TyrValAsnMet-----LeuThrPheThrLeuGluHisGlyGluPhe 152
 QY 686 GAAACGGCTGGGAGAGTCCATATGACCTTAAGCTGCTCAGACATCCCTTTTAATA 745
 DB 153 GluAspGlyLeuGlyLysCysProTyrAspProAlaLysGlyHisAlaGlyLeuVal 172
 QY 746 GATGAGAAATATATCTCTGGAATCGAGCTGATTTTATGGCGGAGACTTCTCTATCTTC 805
 DB 173 AspGlyGluLeuTyrSerAlaThrLeuAsnAsnPheLeuGlyThrGluProIleLeuLeu 192
 QY 806 CGAACTCTTGGCCACCAACCAATCAGGACAGACGATGATTCAGGTGGCTCAAT 865
 DB 193 ArgAsnMetGlyProHisHisSerMetLysThrGlu---TyrLeuAlaPheTrpLeuAsn 211
 QY 866 GATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGACAATCCT-----GAA 916
 DB 212 GluProHisPheValGlySerAlaTyrValProGluSerValGlySerPheThrGlyAsp 231
 QY 917 GATGACAAAGTACTTTTCTTCGTAAGTAAATGCAATAGATGGAGACACTCTCGAAA 976
 DB 232 AspAspLysValTyrPhePhePheArgGluArgAlaValGluSerAspCysTyrAlaGlu 251
 QY 977 GCTACTCAGCTAGATAGGTGATATCAAGATGACTTTGGAGGGCAGCAAGTCTG 1036
 DB 252 GlnValValAlaAlaArgValAlaAlaArgValCysLysGlyAspMetGlyGlyAlaAlaArgThrLeu 271

QY 1037 GTGAATAATCGACAAACATTTCTCAAGCTCGTCTGATTGCTCAGTGCCAGGTCAAAT 1096
 DB 272 GlnArgLysTrpThrThrPheLeuLysAlaArgLeuAlaCysSerAlaPro----- 288
 QY 1097 GGCATTGCACACTCAATTTTGATGAACAGTGCAGGATGTATTCTTAATGAACCTTTAAAGATCCT 1156
 DB 289 AsnTrpGlnLeuTyrPheAsnGlnLeuGlnAlaMetHisThrLeuGlnAspThrSerTrp 308
 QY 1157 AAAAATCCAGTTGATATGAGAGTGTTCAGACTCCAGTAACATTTTCAAGGGATCAGCC 1216
 DB 309 HisAsnThrThrPhePheGlyValPheGlnAlaGlnTrpGlyAspMetTyrLeuSerAla 328
 QY 1217 GTGTGATATGATCAGTGCAGTGCAGAGAGGTGTTCTTGGTCCATATGATGCCACAGG 1276
 DB 329 IleCysGluTyrGlnLeuGluGluIleGlnArgValPheGluGlyProTyrLysGluTyr 348
 QY 1277 GATGACCAACTATCAATGGTGCCTTTATCAAGAAAGAGTCCCTATCCACGGCAGGA 1336
 DB 349 HisGluGluAlaGlnLysTrpAspArgTyrThrAspProValProSerProArgProGly 368
 QY 1337 ACTTGTCACCAACAAACATTT-----GGTGGTTTGTACTCTACAAAGGACCTTCTGAT 1390
 DB 369 SerCysIleAsnAsnTrpHisArgHisGlyTyrThrSerSerLeuGluLeuProAsp 388
 QY 1391 GATCTTAACTTTGCAAGAGTCCATCCAGCCATGTACATCCAGTGTTCCTATGAAC 1450
 DB 389 AsnIleLeuAsnPheValLysLysHisProLeuMetGluGluGlnValGlyProArgTrp 408
 QY 1451 AATGCCCAATAGTGTCAAAACCGATCTAAATTTATCAATTTTACACAAATTTGCTAGAC 1510
 DB 409 SerArgProLeuLeuValLysLysGlyThrAsn-----PheThrHisLeuValAlaAsp 426
 QY 1511 CGATGATGATGAGAGTGA---CAGTATGATGTTATGTTTATCGGAACAGATGTTGGG 1567
 DB 427 ArgValThrGlyLeuAspGlyAlaThrTyrThrValLeuPheIleGlyThrGlyAspGly 446
 QY 1568 ACCGTTCTTAAGTAGTTCAATTTCTTAAGGAGACTTGGTATGATTTTAGAAGAGTTCG 1627
 DB 447 TrpLeuLeuLysAlaValSerLeu-----GlyProTrpValHis-----Leu 460
 QY 1628 CTGGAAGAAATGACAGTTTTT---CGGAAACCGACTGCTATTTTACGAAATGAGAGTTC 1684
 DB 461 IleGluGluLeuGlnLeuPheAspGlnGluPro-----MetArgSerLeuValLeuSer 478
 QY 1685 ACTAAGCAGCAACATATATATATTGTTTCAAGCTGGGTGGCCAGCTCCCTTTTAC 1744
 DB 479 GlnSerLysLysLeuLeuPheAlaGlySerArgSerGlnLeuValGlnLeuProValAla 498
 QY 1745 CGGTGTGATATTTACGGAAAGCGTGTCTGAGTGTGCTCGCCCGAGACCTTACTGT 1804
 DB 499 AspCysMetLysTyr---ArgSerCysAlaAspCysValLeuAlaArgAspProTyrCys 517
 QY 1805 CTTTGGGATGTTCTGCTATGTTCTCGTATTTTCCCACTGCAAAAGAGACGCAACAGCA 1864
 DB 518 AlaTrpSerValAsn-----ThrSerArg 525
 QY 1865 -----CAAGATATTAAGAAATGGAGAC 1885
 DB 526 CysValAlaValGlyGlyHisSerGlySerLeuLeuIleGlnHisValMetThrSerAsp 545
 QY 1886 CCATGCTCACTGTTTCAGACTTACCATCATATACCATGATGATGATGATGATGATGATGAT 1936
 DB 546 ThrSerGlyIleCys-----AsnLeuArgGlySerLysLysValArg 559
 QY 1937 CCTGAAGAGAGAAATCATCTATGTTAGAGAAATAGTACACATTTTGGAAATGCAAGTCCG 1996
 DB 560 ProThrProLysAsnIleThrValValAlaGlyThrAspLeuValLeuProCysHisLeu 579
 QY 1997 AAGTCGCGAGAGCGCTGCTGCTATGCAATTTCCAGAGCGCAAT-----GAAGAGCGA 2050
 DB 580 SerSerAsnLeuAlaHisAlaArgTrpThrPheGlyGlyArgAspLeuProAlaGluGln 599

QY 2051 AAAGAGATCAGAGTGGATGATCATATCATCAGCAGATCAGAGGCTTCTGCTAGT 2110
Db 600 ProGlySerPheLeuTyAspAlaArgLeu-----GlnAlaLeuValMet 615
QY 2111 AGCTTACACAGCAAGATTTCAGGCAATTACCTCTGCATGCGGTGGAACATGGG 2164
Db 616 AlaAlaGlnProArgHisAlaGlyAlaTyHisCysPheSerGluGluGlnGly 633

RESULT 13

US-10-369-072-16
; Sequence 16, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigar, Muralidhara
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 833
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-072-16

Alignment Scores:
Pred. No.: 1.63e-72 Length: 833
Score: 950.50 Matches: 235
Percent Similarity: 50.00% Conservative: 104
Best Local Similarity: 34.66% Mismatches: 262
Query Match: 19.41% Indels: 77
DB: 6 Gaps: 21

US-09-774-490-1 (1-2709) x US-10-369-072-16 (1-833)

QY 206 TGGTTAACTAGGATTGTCTGCTCTTTCTGGGGAGTATTACTTACAGCAAGAGCAAACTAT 265

Db 8 TrpLeu-----LeuAlaAlaArgLeuTrpGlyLeuGlyIleGlyAlaGluValTrpTrp 25
QY 266 CAGAATGGGAAGAAACAATGTCCCAAGGCTGAAATTTATCTTACAAAGAAATGTTGGAATCC 325
Db 26 -----AsnLeuValProArgLysThrValSerSerGlyGlyLeuAlaThrVal 41
QY 326 AACAAATGTGATCACTTTTCATGGCTTGCCCAACAGCTCCAGTTATCATCATCTTCCTTTTG 385
Db 42 ValArgArgPheSerGlnThrGlyLe-----GlnAspPheLeuThrLeuThrLeu 58
QY 386 GATGAGGAACGAGTAGGTGTATGTTGGAGCAAGGATCACATATTTTTCATTCACGCTG 445
Db 59 ThrGluProThrGlyLeuLeuTyValGlyAlaArgAspHisAlaSerAlaLeuGlyVal 78
QY 446 GTTAATATCAAGGATTTTCAAAAGATTGTGTGGCCAGTATCTTACACAGAGAGATGAA 505
Db 79 ProValLeuLeuLeuGlnAlaValIleSerTrpGluAlaProValGluLysThrGlu 98
QY 506 TGCAAGTGGGCTGGAAGACATCTGAAAGAATGTGCTAAATTTTCATCAAGGTACTTAAG 565
Db 99 CysIleGlnLysGlyLysAsnAsnGlnThrGluCysPheAsnPheIleArgPheLeuGln 118
QY 566 GCATATAATCACTCACTTGTACGCTGTGGAACGGGGGCTTTTCATCAATTTGCACC 625
Db 119 ProTyAsnAlaSerHisLeuTyValCysGlyThrTyAlaPheGlnProLysCysThr 138
QY 626 TACATTGAAATTTGGACATCATCTCGAGGACAATATTTTAAAGCTGGAGCACTCACATTTT 685
Db 139 TyrValAsnMet-----LeuThrPheThrLeuGluHisGlyGluPhe 152
QY 686 GAAACCGCGCTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGCATCCCTTTTAATA 745
Db 153 GluAspGlyLysGlyLysCysProTyAspProAlaLysGlyHisAlaGlyLeuLeuVal 172
QY 746 GATGAGAAATTAATCTGTTGAACTGCAGCTGATTTTATGGGCGAGACTTGTCTATCTTC 805
Db 173 AspGlyGluLeuTySerAlaThrLeuAsnAsnPheLeuGlyThrGluProIleLeu 192
QY 806 CGAACTCTTGGGCACCAACCACCTCATCTCAGAGAGTGACAATCTCT-----GAA 916
Db 212 GluProHisPheValGlySerAlaTyValProGluSerValGlySerPheThrGlyAsp 231
QY 917 GATGACAAAGATATACTTTTCTTCGTGAAATGCAATAGATGAGAGAACACTTGGAAAA 976
Db 232 AspAspLysValTyPhePheArgGluAlaValGluSerAspCysTyAlaGlu 251
QY 977 GCTACTCAGCTAGATAGGTACATATGCAAGTATGCTTTGGAGGCACAGAGTCTG 1036
Db 252 GlnValValAlaArgValAlaArgValCysLysGlyAspMetGlyGlyAlaArgThrLeu 271
QY 1037 GTGAATAATGGACAACATTTCTCAAGCTCGTCTGATTTCTCAGTCCAGGTCCTCAAT 1096
Db 272 GlnArgLysTrpThrThrPheLeuLysAlaArgLeuAlaCysSerAlaPro----- 288
QY 1097 GGCATTGACACTCAATTTTGATGAAGTGTATTCCTTAATGAACCTTTAAAGATCCT 1156
Db 289 AsnTrpGlnLeuTyPheAsnGlnLeuGlnAlaMetHisThrLeuGlnAspThrSerTrp 308
QY 1157 AAAAATCCAGTGTATATGGAGTGTTCACGACTTCCAGTAAACATTTTCAGGGATCAGCC 1216
Db 309 HisAsnThrThrPhePheGlyValPheGlnAlaGlnTrpGlyAspMetTyLeuSerAla 328
QY 1217 GTGTGTATGATAGCATGATGATGAGAGGGGTTCCTTGTGCTCATATGCCACAGG 1276
Db 329 IleCysGluTyGlnLeuGluGluIleGluArgValPheGluGlyProTyLysGluTy 348
QY 1277 GATGGACCAACTATCAATCGGTGCTTTATCAAGGAAGAGTCCCTTATCCACGCCACGA 1336

```
Db 349 HisGluGluAlaGlnLysTrpAspArgTyrThrAspProValProSerProArgProGly 368
Qy 1337 ACTTGTCGCCAGCAAAACATTT-----GGTGGTTTGTACTCTACAAAGAGACTTCTCAT 1390
Db 369 SerCysIleAsnAsnTrpHisArgHisGlyTyrThrSerSerLeuGluLeuProAsp 388
Qy 1391 GATGTTTAAACCTTTGCGAAGAGTCATCCAGCAGTCATCAATCCAGTGTTCCTCATGAAC 1450
Db 389 AsnIleLeuAsnPheValLysLysHisProLeuMetGluGluGlnValGlyProArgTrp 408
Qy 1451 AATGCCCAATAGTGCATCAAAACGGATGTAATATCAATTATACAAATATGTCGTAGAC 1510
Db 409 SerArgProLeuLeuValLysLysGlyThrAsn-----PheThrHisLeuValAlaAsp 426
Qy 1511 CGAGTGGATCGAAGATGGA---CAGTATGATGTATGTTTATCGGAACAGATGTTGGG 1567
Db 427 ArgValThrGlyLeuAspGlyAlaThrTyrThrValLeuPheIleGlyThrGlyAspGly 446
Qy 1568 ACCGTTCTTAAAGTAGTTTCAATTCCTAAGGAGACTTGTGTATGATTTAGAAAGAGTTCTG 1627
Db 447 TrpLeuLeuLysAlaValSerLeu-----GlyProTrpValHis-----Leu 460
Qy 1628 CTGGAAGAAATGACAGTGT---CGGAACCGACTGCTATTATTCAGCAATGAGCTTTCC 1684
Db 461 IleGluGluLeuGlnLeuPheAspGlnPro-----MetArgSerLeuValSer 478
Qy 1685 ACTAGCAGCAACAATATATATTGTTCAACGGCTGGGTTGCCAGCTCCCTTTACAC 1744
Db 479 GlnSerLysLeuLeuPheAlaGlySerArgSerGlnLeuValGlnLeuProValAla 498
Qy 1745 CGGTGTGATATTACGGAAACGGTGTCTGAGTGTGCTCGCCGAGACCTTACTGT 1804
Db 499 AspCysMetLysTyr---ArgSerCysAlaAspCysValLeuAlaIleAspProTyrCys 517
Qy 1805 GCTTCGGATGTTTCGATGTTCTCGCTATTTTCCACTGCAAGAGAGACGACAGAGA 1864
Db 518 AlaTrpSerValAsn-----ThrSerArg 525
Qy 1865 -----CAAGATATAAGAAATGGAGAC 1885
Db 526 CysValAlaValGlyGlyHisSerGlySerLeuLeuIleGlnHisValMetThrSerAsp 545
Qy 1886 CCACGTGACTCAGTTCAGACTTACACCATGATATACCATGTCAGGCGCACAGC----- 1936
Db 546 ThrSerGlyIleCys-----AsnLeuArgGlySerLysLysValArg 559
Qy 1937 CTTGAAGAGAGAAATCATCTATGCTGATAGATAGTACACATTTTTCGAATGCGAGTCCG 1996
Db 560 ProThrProLysAsnIleThrValValAlaGlyThrAspLeuValLeuProCysHisLeu 579
Qy 1997 AAGTCGACAGAGCGCTGCTATTGGCAATTCACAGAGCGGAAAT-----GAAGAGCGA 2050
Db 580 SerSerAsnLeuAlaHisAlaArgTrpThrPheGlyArgAspLeuProAlaGln 599
Qy 2051 AAAGAAGAGATCAGAGTGATCATATCATCATCAGACAGATCAAGGCTTCTGCTACGT 2110
Db 600 ProGlySerPheLeuTyrAspAlaArgLeu-----GlnAlaLeuValValMet 615
Qy 2111 AGCTTACACAGAGAGATTCAGGCAATTACCTCTCCCATGCGGTGGAAATCGG 2164
Db 616 AlaAlaGlnProArgHisAlaGlyAlaTyrHisCysPheSerGluGluGlnGly 633

RESULT 14
US-10-369-072-18
; Sequence 18, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
```

```
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Iaca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shinkets, Richard
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 833
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-072-18
```

Alignment Scores:

Pred. No.:	1,63e-72	Length:	833
Score:	950.50	Matches:	235
Percent Similarity:	50.00%	Conservative:	104
Best Local Similarity:	34.66%	Mismatches:	262
Query Match:	19.41%	Indels:	77
DB:	6	Gaps:	21

US-09-774-490-1 (1-2709) x US-10-369-072-18 (1-833)

```
Qy 206 TGGTTAACTAGGATGTGTCTTTCTGGGGAGTATTACTTACAGCAAGAGCAAACTAT 265
Db 8 TrpLeu-----LeuAlaAlaArgLeuTrpGlyLeuGlyIleGlyAlaGluValTrp 25
Qy 266 CAGAATCGGAAGAACAATGTCACAGGCTGAAATATCTCAAGAAGAAATGTTGGAATCC 325
Db 26 -----AsnLeuValProArgLysThrValSerSerGlyGlyLeuAlaThrVal 41
Qy 326 AACAAATGTGATCACTTTTCAATGGCTTGGCCCAACAGCTCCAGCTTATCATCTTCTTTTG 385
Db 42 ValArgArgPheSerGlnThrGlyIle-----GlnAspPheLeuThrLeu 58
Qy 386 GATCAGGAACGAGTACGCTGTATGTTGGAGCAAGAGTACATATTTTTCATTCGACCTG 445
Db 59 ThrGluProThrGlyLeuLeuTyrValGlyAlaArgAspHisAlaSerAlaLeuGlyVal 78
Qy 446 GTTAATATCAAGGATTTTCAAAGATTTGTGGCCAGTATCTTACACCAAGAGAGATCAA 505
Db 79 ProValLeuLeuLeuGlnAlaValIleSerTrpGluAlaProValGluLysThrGlu 98
```

```

QY 506 TCGAAGTGGCTGGAAGACATCTCTGAAGATGCTAATTTTCATCAAGTACTTAAG 565
Db 99 CysIleGlnLysGlyLeuAsnGlnThrGluCysPheAsnPhelIleArgPheLeuGln 118
QY 566 GCATATAATCAGACTCAGCTTGTACGCTGTGGAACGGGGCTTTTCATCCAAATTTGCACC 625
Db 119 ProTyrAsnAlaSerHisLeuTyrValCysGlyThrTyrAlaPheGlnProLysCysThr 138
QY 626 TACATTGAAATTGGACATCATCTCTGAGGACAAATATTTTAAAGCTGGAGAACTCACAATTT 685
Db 139 TyrValAsnMet-----LeuThrPheThrLeuGluHisGlyGluPhe 152
QY 686 GAAACGGCGCTGGGAAGAGTCCATATGACCTTAAGCTGCTCAGACGATCCCTTTTATA 745
Db 153 GluAspGlyLysGlyLeuCysProTyrAspProAlaLysGlyHisAlaGlyLeuVal 172
QY 746 GATGAGAAATATATCTCTGGAACCTCAGCTGATTTTATGGGCGAGACTTTGCTATCTTC 805
Db 173 AspGlyGluLeuTyrSerAlaThrLeuAsnAsnPheLeuGlyThrGluProIleLeu 192
QY 806 CGAATCTTGGGACACCAATCAGACAGACGATGATTCAGAGTGGCTCAAT 865
Db 193 ArgAsnMetGlyProHisHisSerMetLysThrGlu---TyrLeuAlaPheThrLeuAsn 211
QY 866 GATCCAAAGTTCAATTAGTGGCCACCTCATCTCAGAGAGTGACATCCT-----GAA 916
Db 212 GluProHisPheValGlySerAlaTyrValProGluSerValGlySerPheThrGlyAsp 231
QY 917 GATGCAAGTATATCTTTCTTCGTCGAAATGCAATAGATGAGAGACACTCTGGAAGA 976
Db 232 AsplepLysValTyrPhePhePheArgGluArgAlaValGluSerAspCysTyrAlaGlu 251
QY 977 GCTACTACGCTAGATAGTGCAGATATGCAAGATATGCTTTGGAGGCGACAGAGTCTG 1036
Db 252 GlnValValAlaArgValAlaArgValCysLysGlyAspMetGlyGlyAlaArgThrLeu 271
QY 1037 GTGAATAATGACACATCTCTCAAGCTCGTGTGATTTGCTCAGTCCAGGTCCTCAAT 1096
Db 272 GlnArgLysTyrThrPheLeuLysAlaArgLeuAlaCysSerAlaPro----- 288
QY 1097 GGCATTGACACTCATTTGATGAATGCTGAGATGATTTCTTAATGAATCTTAAAGATCCT 1156
Db 289 AsnTrpGlnLeuTyrPheAsnGlnLeuGlnAlaMetHisThrLeuGlnAspThrSerTrp 308
QY 1157 AAAATCCAGTTGATATGAGTGTATACGACTTCCAGTAAACATTTTCAAGGATCAGCC 1216
Db 309 HisAsnThrThrPhePheGlyValPheGlnAlaGlnTrpGlyAspMetTyrLeuSerAla 328
QY 1217 GTGTGTATATAGCATGATGATGTGAGAGGTGTTCTTGTGCTCATATGCCACAGG 1276
Db 329 IleCysGluTyrGlnLeuGluGluIleGlnArgValPheGluGlyProTyrLysGluTyr 348
QY 1277 GATGACCAACTATCATATGCTGCTTATCAAGGAGAGTCCCTATCCACGCGCCAGGA 1336
Db 349 HisGluGluAlaGlnLysTrpAspArgTyrThrAspProValProSerProArgProGly 368
QY 1337 ACTTGTCACCAAAACATTT-----GGTGGTTTGTACTTCAAAAGACCTTCTCGAT 1390
Db 369 SerCysIleAsnAsnTrpHisArgArgHisGlyTyrThrSerSerLeuGluLeuProAsp 388
QY 1391 GATGTTAACTTTGCAAGAAGTATCCAGCCATGATCAATTCAGTGTTCCTATGAGAC 1450
Db 389 AsnIleLeuAsnPheValLysLysHisProLeuMetGluGluGlnValGlyProArgTrp 408
QY 1451 AATGCCCAATAGTATCAAAACGGATGTAATATCAATTTACAAATTTGCTGTAGAC 1510
Db 409 SerArgProLeuLeuValLysLysGlyThrAsn-----PheThrHisLeuValAlaAsp 426
QY 1511 CGAGTGGATGAGAGAGTGA---CAGTATGATGTTATGTTTATCGGAACAGAGTGTGGG 1567
Db 427 ArgValThrGlyLeuAspGlyAlaThrTyrThrValLeuPheIleGlyThrGlyAspGly 446
QY 1568 ACCGTTCTTAAAGTAGTTTCAATCTTCAATGAGAGACTTGGTATGATTTAGAGAGGTTCTG 1627

```

```

Db 447 TrpLeuLeuLysAlaValSerLeu-----GlyProTrpValHis-----Leu 460
QY 1628 CTGGAAGAAATGACAGATTTT---CGGAACCGACTGCTATTTTCAGCAATGAGCTTTCC 1684
Db 461 IleGluGluLeuGlnLeuPheAspGlnIlePro-----MetArgSerLeuValLeuSer 478
QY 1685 ACTAAGCAGCAACAATATATATTTGGTTCAACGGCTGGGGTGGCCAGCTCCCTTTACAC 1744
Db 479 GlnSerLysLysLeuLeuPheAlaGlySerArgSerGlnLeuValGlnLeuProValAla 498
QY 1745 CGGTGTGATATTTACGGGAAGCGTGTCTCAGTGTTCCTCGCCGCGAGACCTTACTGT 1804
Db 499 AspCysMetLysTyr---ArgSerCysAlaAspCysValLeuAlaArgAspProTyrCys 517
QY 1805 GCTTGGGATGTTCTGCATGTTCTCGCTATTTTCCACTGCAAAAGAGAGCACAAGACGA 1864
Db 518 AlaTrpSerValAsn-----ThSerArg 525
QY 1865 -----CAAGATATAGAATAATGGAGAC 1885
Db 526 CysValAlaValGlyGlyHisSerGlySerLeuLeuIleGlnHisValMetThrSerAsp 545
QY 1886 CCATGCTACTCTGTTTACAGCTTACACCATGATATACCATGCGCCACAGC----- 1936
Db 546 ThrSerGlyIleCys-----AsnLeuArgGlySerLysLysValArg 559
QY 1937 CCTCAAGAGAGAAATCATCTATGTTGTAGAGAATAGTACACATTTTGGAAATGAGTCCG 1996
Db 560 ProThrProLysAsnIleThrValValAlaGlyThrAspLeuValLeuProCysHisLeu 579
QY 1997 AAGTCGACAGAGCGCTGCTTATTTGGCAATTCAGAGCGCAAT-----GAAGAGCGA 2050
Db 580 SerSerAsnLeuAlaHisAlaArgTrpThrPheGlyGlyArgAspLeuProAlaGluGln 599
QY 2051 AAAGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGCT 2110
Db 600 ProGlySerPheLeuTyrAspAlaArgLeu-----GlnAlaLeuValValMet 615
QY 2111 AGTCTACAACAGAGGATTCAGGCAATTTACCTCTGCCATTCGCGTGGGAACATGGG 2164
Db 616 AlaAlaGlnProArgHisAlaGlyAlaTyrHisCysPheSerGluGluGlnGly 633

RESULT 15
PCT-US03-09929-2
; Sequence 2, Application PC/JUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.

```

; NUMBER OF SEQ ID NOS: 179
 ; SOFTWARE: CuraseqList version 0.1
 ; SEQ ID NO 2
 ; LENGTH: 634
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 PCT-US03-09929-2

Alignment Scores:

Pred. No.: 1.82e-72 Length: 634
 Score: 949.50 Matches: 230
 Percent Similarity: 50.76% Conservative: 102
 Best Local Similarity: 35.17% Mismatches: 251
 Query Match: 19.39% Indels: 71
 DB: 1 Gaps: 19

US-09-774-490-1 (1-2709) x PCT-US03-09929-2 (1-634)

QY 278 AACAAATGTCAGAGCTGAAATATCTTACAAAGAAATGTTGGAATCCAAACAATGTGATC 337
 DB 6 AsnLeuValProArgLysThrValSerSerGlyGluLeuAlaThrValValArgPhe 25
 QY 338 ACTTTCATGCTGGCCCAACAGCTCCAGTTATCATACCTCTCTTTGGATGAGGAACGG 397
 DB 26 SerGlnThrGlyLeu-----GlnAspPheLeuThrLeuThrGluProThr 42
 QY 338 AGTAGCTGTATGTTGGAGCAAGGATCACATATTTTCATTCGACTGGTTAATATCAAG 457
 DB 43 GlyLeuLeuTyValGlyAlaArgGluAlaLeuPheAlaPheSerMetGluAlaLeuGlu 62
 QY 458 GATTTCAAAAGATTGTGGCCAGTACTTACACAGAGAGATGAATGCAAGTGGGCT 517
 DB 63 LeuGlnGlyAlaIleSerTrpGluAlaProValGluGluLysThrGluCysIleGlnLys 82
 QY 518 GGAAAGACATCCTGAAAGAAATGTCTAATTTTCATCAAGGTACTTTAAGGCATATAATCAG 577
 DB 83 GlyLysAsnAsnGlnThrGluCysPheAsnPhelIleArgPheLeuGlnProTrpAsnAla 102
 QY 578 ACTGACTGTAGCCCTGTGGAGCGGGGCTTTTCATCCAAATTTGACCTACATCAATTAAT 637
 DB 103 SerHisLeuTyValCysGlyThrTyAlaPheGlnProLysCysThrTyValAsnMet 122
 QY 638 GGACATCATCCTGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTTGAAGACGGCGT 697
 DB 123 -----LeuThrPheThrLeuGluHisGlyGluPheGluAspGlyLys 136
 QY 698 GGGAGAGTCCATATGACCTAAGCTGTGACAGCATCCCTTTTAAATAGATGGAGAATTA 757
 DB 137 GlyLysCysProTyAspProAlaLysGlyHisAlaGlyLeuLeuValAspGlyGluLeu 156
 QY 758 TACTCTGGACGTGAGCTGATTTTATGGGCGAGACTTTGCTATCTTCGAACTCTTGGG 817
 DB 157 TyrSerAlaThrLeuAsnAsnPheLeuGlyThrGluProIleIleLeuArgAsnMetGly 176
 QY 818 CACCACCACTCCATGAGCAGCAGCATGATTCAGGTGCTCAATGATCCAAAGTTC 877
 DB 177 ProHisHisSerMetCysThrGlu---TyrLeuAlaPheTrpLeuAsnGluProHisPhe 195
 QY 878 ATTAGTGCCCACTCATCTCAGAGAGTGACAAATCCT-----GAAGATGACAAAGTA 928
 DB 196 ValGlySerAlaTyValProGluSerValGlySerPheThrGlyAspAspLysVal 215
 QY 929 TACTTTTCTTCGTTGAAATCAATAGATGGAGAACACTCTGGAAAAGCTACTACCGCT 988
 DB 216 TyrPhePhePheArgGluArgAlaValGluSerAspCysTyAlaGluGlnValValAla 235
 QY 989 AGAATAGGTGCATATGCAAGAAATCACTTTTGAGGCGCACAGAAAGTCTGTGTAATAATGG 1048
 DB 236 ArgValAlaArgValCysLysGlyAspMetGlyGlyAlaArgThrLeuGlnArgLysTrp 255
 QY 1049 ACAACATCTCTCAAGCTGCTGATTTGCTCAGTGCAGGCTCCAAATGGCATTCACACT 1108
 DB 255 ThrThrPheLeuLysAlaArgLeuAlaCysSerAlaPro-----AsnTrpGlnLeu 272

QY 1109 CATTTTCATGAAGCTGAGGATGTTATCTCTAATGAATTTTAAAGATCTCTAAATCCAGTT 1168
 DB 273 TyrPheAsnGlnLeuGlnAlaMetHisThrLeuGlnAspThrSerTrpHisAsnThrThr 292
 QY 1169 GTATATGAGAGTGTTCACGACTTCCAGTAACATTTTCAAGGGATCAGCGGTGTGTATGAT 1228
 DB 293 PhePheGlyValPheGlnAlaGlnTrpGlyAspMetTyrLeuSerAlaIleCysGluLys 312
 QY 1229 AGCATAGTATGTGAGAGGGTGTCTTCTGTTGTCATATGCCACAGGATGGACCCCAAC 1288
 DB 313 GlnLeuGluGluIleGlnArgValPheGluGlyProTyrLysGluTyHisGluGluAla 332
 QY 1289 TATCAATGGGTGCTTATCAAGGAAGAGTCCCTATCCAGCGCCAGGAACCTTGTCCCAAGC 1348
 DB 333 GlnLysTrpAspArgTyrThrAspProValProSerProArgProGlySerCysIleAsn 352
 QY 1349 AAAACATTT-----GGTGGTTTGTACTCTACAAAGGACCTTCTCTGATGATGTTATACC 1402
 DB 353 AsnTrpHisArgArgHisGlyTyrThrSerSerLeuGluLeuProAspAsnIleLeuAsn 372
 QY 1403 TTTCGAAGAAGTGTATCCAGCCATGTACATCCAGTGTCTCTATGAACAATCGCCCAATA 1462
 DB 373 PheValLysLysHisProLeuMetGluGluGlnValGlyProArgTrpSerArgProLeu 392
 QY 1463 GTGATCAAAACGGATGTAAATTTATCAATTTTACACAAATTTGCTAGACCGAGTGGATGCA 1522
 DB 393 LeuValLysLysGlyThrAsn-----PheThrHisLeuValAlaAspArgValThrGly 410
 QY 1523 GAAGATCGA---CAGTATGATGTTATGTTTATCGGAACAGATGTTGGACCGTCTTTAAA 1579
 DB 411 LeuAspGlyAlaThrTyThrValLeuPheIleGlyThrGlyAspGlyTrpLeuLys 430
 QY 1580 GTAGTTTCAATCTTAAGGAGACTTGTGTATGATTAGAGAGGTTCTGCTCGAAGAAATG 1639
 DB 431 AlaValSerLeu-----GlyProTrpValHis-----LeuIleGluGluLeu 444
 QY 1640 ACAGTTT---CGGGAACCGACTGCTATTTTCAGCAATGGAGCTTCCACTAAGCAGCAA 1696
 DB 445 GlnLeuPheAspGlnGluPro-----MetArgSerLeuValLeuSerGlnSerLys 462
 QY 1697 CAATATATATGCTTCAACGGCTGGGTGTCGCCAGCTCCCTTTACACCGGTGTGATATT 1756
 DB 463 LeuLeuPheAlaGlySerArgSerGlnLeuValGlnLeuProValAlaAspCysMetLys 482
 QY 1757 TACGGGAAGCGTGTGTGAGTGTGCTGCTGCCGAGACCCCTTACTGCTGCTGGATGGT 1816
 DB 483 Tyr---ArgSerCysAlaAspCysValLeuAlaArgAspProTyrCysAlaTrpSerVal 501
 QY 1817 TCTGCATGTTCTCGCTATTTTCCCACTGCAAGAGAGCGCAACAGCA-----1864
 DB 502 Asn-----ThrSerArgCysValAlaVal 509
 QY 1865 -----CAAGATATAAGAAATGGAGACCCCACTGACTCAC 1897
 DB 510 GlyGlyHisSerGlySerLeuLeuIleGlnHisValMetThrSerAspThrSerGlyLe 529
 QY 1898 TGTTTCAGACTTACACCATGATAATCACCATGGCCACAGC-----CCTGAAGAGAGA 1948
 DB 530 Cys-----AsnLeuArgGlySerLysValArgProThrProLys 543
 QY 1949 ATCATCTATGTTAGAGATAGTAGCACATTTTGGAAATGTCAGTCCGAAGTCCGACAGA 2008
 DB 544 AsnIleThrValValAlaGlyThrAspLeuValLeuProCysHisLeuSerSerAsnLeu 563
 QY 2009 GCGTGTGCTTATTCGCAATTCACAGAGCGCAAT-----GAAGAGCGAAAGAAGAGATC 2062
 DB 564 AlaHisAlaArgTrpThrPheGlyArgAspLeuProAlaGlnGlnProGlySerPhe 583
 QY 2063 AGAGTGGATGATCATATCATCAGCAGATCAAGGCCTTCTGCTACGTAGTCTACAACAG 2122
 DB 584 LeuTyrAspAlaArgLeu-----GlnAlaLeuValValMetAlaAlaGlnPro 599

Qy 2123 AAGGATTGAGCAATTACCTCTGCCATGCGGTGGGAACATGGG 2164
::: :||| ||| ||| ::: ||| |||
Db 600 ArgHisAlaGlyAlaTyrHisCysPheSerGluGluGlnGly 613

Search completed: August 3, 2003, 10:21:13
Job time : 44 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 09:57:04 ; Search time 27 Seconds
(without alignments)
8490.379 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 4896

Sequence: 1 aattttttatttcatgatg.....agggtttttttctctaataacc 2709

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 657434

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-O=/cgn2_1/USPTO spool/US09774490/runat 03082003 095416 25355/app query.fasta_1.2887
-DB=Issued Patents AA -QFMT=fastan -SUFFIX=xsi -MINMATCH=0.1 -IOFCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=prco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09774490 @CGN 1 1 30 @runat 03082003 095416 25355 -NCPUS=6 -ICPU=3
-NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents AA:
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pdp:*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pdp:*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pdp:*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pdp:*
5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pdp:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pdp:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4201	85.8	771	1	US-08-121-713D-54
2	4201	85.8	771	1	US-08-835-268-54
3	4201	85.8	771	2	US-09-060-692-54
4	4201	85.8	771	3	US-08-833-391-54
5	4201	85.8	771	4	US-09-060-610-54
6	4201	85.8	771	5	PCT-US94-10151A-54
7	3568	72.9	655	4	US-08-556-422A-3
8	2562	52.3	477	1	US-08-136-922-2
9	2030	41.5	775	4	US-09-308-179B-1
10	982.5	20.1	862	4	US-08-556-422A-2
11	838.5	17.1	607	4	US-08-556-422A-4
12	785.5	16.0	724	1	US-08-121-713D-62

13 785.5 16.0 724 1 US-08-835-268-62 Sequence 62, Appl
14 785.5 16.0 724 2 US-09-060-692-62 Sequence 62, Appl
15 785.5 16.0 724 3 US-08-833-391-62 Sequence 62, Appl
16 785.5 16.0 724 4 US-09-060-610-62 Sequence 62, Appl
17 785.5 16.0 724 5 PCT-US94-10151A-62 Sequence 62, Appl
18 785.5 15.6 930 4 US-09-254-594-6 Sequence 6, Appl
19 762.503 15.6 730 1 US-08-121-713D-58 Sequence 58, Appl
20 762.5 15.6 730 1 US-08-835-268-58 Sequence 58, Appl
21 762.5 15.6 730 2 US-09-060-692-58 Sequence 58, Appl
22 762.5 15.6 730 3 US-08-833-391-58 Sequence 58, Appl
23 762.5 15.6 730 4 US-09-060-610-58 Sequence 58, Appl
24 762.5 15.6 730 5 PCT-US94-10151A-58 Sequence 58, Appl
25 738.503 15.1 650 1 US-08-121-713D-60 Sequence 60, Appl
26 738.5 15.1 650 2 US-08-835-268-60 Sequence 60, Appl
27 738.5 15.1 650 2 US-09-060-692-60 Sequence 60, Appl
28 738.5 15.1 650 3 US-08-833-391-60 Sequence 60, Appl
29 738.5 15.1 650 4 US-09-060-610-60 Sequence 60, Appl
30 738.5 15.1 650 5 PCT-US94-10151A-60 Sequence 60, Appl
31 723.5 14.8 929 4 US-09-254-594-3 Sequence 3, Appl
32 723 14.8 887 4 US-09-077-940A-2 Sequence 2, Appl
33 718 14.7 425 4 US-08-556-422A-7 Sequence 7, Appl
34 712.5 14.6 888 4 US-09-077-940A-4 Sequence 4, Appl
35 680.003 13.9 712 1 US-08-121-713D-64 Sequence 64, Appl
36 680 13.9 712 1 US-08-835-268-64 Sequence 64, Appl
37 680 13.9 712 2 US-09-060-692-64 Sequence 64, Appl
38 680 13.9 712 3 US-08-833-391-64 Sequence 64, Appl
39 680 13.9 712 4 US-09-060-610-64 Sequence 64, Appl
40 680 13.9 712 5 PCT-US94-10151A-64 Sequence 64, Appl
41 654.5 13.4 666 3 US-09-240-410-2 Sequence 2, Appl
42 649.5 13.3 634 3 US-09-041-236-2 Sequence 2, Appl
43 649.5 13.3 634 4 US-09-771-467C-2 Sequence 2, Appl
44 598.5 12.2 295 4 US-08-556-422A-6 Sequence 6, Appl
45 534 10.9 606 3 US-09-041-236-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-08-121-713D-54
; Sequence 54, Application US/08121713D
; Patent No. 5639856
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713D
; FILING DATE: 13-SEP-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:

; INFORMATION FOR SEQ ID NO: 54:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 771 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-121-713D-54

Alignment Scores:

Pred. No.: 0 Length: 771

Score: 4201.00 Matches: 771

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 85.80% Indels: 0

DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-121-713D-54 (1-771)

QY 200 ATGGCTGGTTAACTAGGATTCGTCTTTCTGGGAGTATTACTTACAGCAAGACA 259

DB 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20

QY 260 AACTATCAGAAATGGGAAGAACAAATGTGCCAGCTCAAAATTATCTCAAAAGAAATGTG 319

DB 21 AsnTyrGlnAsnGlyLysAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40

QY 320 GAATCAACAATGTGATCACTTTCAATGGCTTGGCCACAGCTCCAGTTATCATCTTC 379

DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60

QY 380 CTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAGAGATCACATATTTTCATTC 439

DB 61 LeuLeuAspGluLysArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80

QY 440 GACCTGGTTAATATCAAGATTTTCAAAAGATTGTGTGCCAGTATCTTACACCAAGA 499

DB 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100

QY 500 GATGAATGCAATGGCTGGAAAGACATCTCGAAGATGTGCTAATTTTCATCAAGTA 559

DB 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120

QY 560 CTTAAGGCATATAATCAGACTCACCTGTACGCTGTGGAACGGGGCTTTTCATCCAATT 619

DB 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140

QY 620 TCACCTACATTTGAAATGGACATCATCTCGAGCAATATTTTAACTGGAGAACTCA 679

DB 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160

QY 680 GATTTTGAACGGCTGGGAAGAGTCCATATGACCCCTAAGCTCTGCACAGCATCCCTT 739

DB 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180

QY 740 TTAATAGATGGAGATTACTCTGGAATCGAGCTGATTTTATGGGGAGACTTTGCT 799

DB 181 LeuIleAspGlyValLeuTyrSerGlyThrAlaIleAspPheMetGlyArgAspPheAla 200

QY 800 ATCTTCCGAACCTTTGGGCACCACCAATCAGGACAGCAGCAGCATGATTCAGGTGG 859

DB 201 IlePheArgThrLeuGlyHisHisHisProIleArgThrGluGlnHisAspSerArgTrp 220

QY 860 CTCATGATCCAAAGTTCTTAGTCCACCTCATCTCAGAGCTGACATCTCTCAAGAT 919

DB 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240

QY 920 GACAAAGTATCTTTTCTCGTGAAATGCAATAGATGGAGAACACTCTCGAAAGCT 979

DB 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyValHisSerGlyLysAla 260

QY 980 ACTCAGCTAGATAGGTGAGATATGCAAGATGATTTGGAGGGCAGCAAGCTGTGGT 1039

DB 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280

QY 1040 AATAATGGCAACATTCCTCAAAGCTGCTGATTGCTCAGTCCAGGTCCTCAATGGC 1099

DB 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300

QY 1100 ATTGACACTCATTTTGTGAACTGCAGGATGTTTCTTAATGAATCTTAAAGATCCTAAA 1159

DB 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320

QY 1160 AATCCAGTTGTATATGAGTGTTCAGACTTCCAGTAACATTTTCAAGGGATCAGCCGTG 1219

DB 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340

QY 1220 TGTATGTATAGCATGATGTGAGAAGGGTGTTCCTTGTGTCCATATGCCACACAGGAT 1279

DB 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360

QY 1280 GGACCCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCACGCCAGGAAT 1339

DB 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380

QY 1340 TGTCCGAGCAAAACATTTGGTGGTTTGTACTCTCAAGAGACCTTCTCTGATGATGTTATA 1399

DB 381 CysProSerLysThrPheGlyPheAspSerThrLysAspLeuProAspAspValIle 400

QY 1400 ACCTTTGCAAGAACTCATCCAGCCATGTACAATCCAGTGTTCCTATGAAACAATCGCCA 1459

DB 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420

QY 1460 ATAGTGTCAAAACCGGATGTAATTTATCAATTTTACAAATTTGTCTGAGACCGAGTGAAT 1519

DB 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440

QY 1520 GCAGAAGATGCACAGTATGATTTATGTTTATCCGAACAGATGTGGACCGTCTTTAAA 1579

DB 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460

QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGTGTATGATTTAGAAAGAGTTCCTCTCGAAGAAATG 1639

DB 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluGluMet 480

QY 1640 ACAGTTTTTCGGGAACCGACTGTATTTTCAGCAATGGAGCTTTCACCTAAGCAGCAACA 1699

DB 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500

QY 1700 CTATATATGTTTCAAAGCTGGGTGCCAGCTCCCTTTACACCGGTGTGATATTTAC 1759

DB 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520

QY 1760 GGGAAAGCTGCTGCTGAGTGTTCCTCGCCGAGACCCCTTACTGTCTTGGGATGGTTCCT 1819

DB 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540

QY 1820 GCATGTTCTCCTATTTTCCCACTGCCAAGAGACGCAACAGCAGCAAGATATAAGAAAT 1879

DB 541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn 560

QY 1880 GGAGACCCACTGACTCTGTTTACAGCTTACCATGATATCACCATGGCCACAGCCCT 1939

DB 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580

QY 1940 GAAGAGAAATCATCTATGTTAGATAATAGTAGACATTTTGGAAATGCAGTCCGGAAG 1999

DB 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600

QY 2000 TCGCAGAGAGCGCTGGTCTATTGGCAATTCAGAGCGGAAATGAAGACGCAAGAAAGAG 2059

DB 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620

QY 2060 ATCAGATGATGATCATATCATCAGGACAGATCAAGCCCTTCTGCTACGTAGTCTACAA 2119

DB 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640

QY 2120 CAGAGGATTGAGCAATTAACCTCTGCATCGGTGGACATGGGTTCATACAAACTCTT 2179
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAAACCTCGAAGTCAATGACACAGAGCAATTTGGAGAACTTCTTCATAAAGAT 2239
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
QY 2240 GATGATGAGATGGCTCTTAAGACCAAGAAATGTCCAATAGCATGACACACCTAGCCAGAAG 2299
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGGTACAGAGCTTTCATGAGCTCATCAACACCCCAATCTCAACACAGATGATGAG 2359
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAGTTTGGAAAGGACCGAAACCAACCTCGCAAGGCCAGGACATACC 2419
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CAGGGAACAGTAAACAAATGGAAGCACTTACAGAAATAAGAAAGGTAGAAACAGAGG 2479
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCACGAAATTTGAGAGGCCACCCAGGAGTGC 2512
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 2

US-08-835-268-54
; Sequence 54, Application US/08835268
; Patent No. 5807826
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,268
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein

US-08-835-268-54

Alignment Scores:

Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-835-268-54 (1-771)

QY 200 ATGGGCTGGTTAACTAGAGATTGTCTCTTTTGGGGAGTATTACTTTACAGCAAGAGCA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGATGGGAAGAACAAATGTCCCAAGGCTGAATTTATCTCTCAAGAAATCTTG 319
Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAACAATGTGATCACTTTCAATGGCTTGGCCCAACAGCTCCAGTTATCATACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyrHisThrPhe 60
QY 380 CTTTGGATGAGGAACCGAGTAGGCTGTATGTTGGAGCAAGAGTACATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAATATCAAGGATTTTCAAAGATTGTGTGGCCAGTATCTTACACCAAGAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATGATGCAAGTGGGCTGGAAAGACATCTCTGAAAGATGTGCTAAATTTTCATCAAGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCTTGTACGCCCTGTGGAACGGGGCTTTTCATCCAATT 619
Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATGAAATTTGACATCATCTTGAGAGCAATATTTTAGCTGGAGAACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTCAAAACGGCCGTGGAGAGTCCATATCACCTTAAGCTGCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGAGAATTATATCTCTGGAACCTCCAGCTCAATTTATGGGGCGAGACTTTGCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCGAACTCTTGGGCAACCAACCCCAATCAGGACAGCAGCATGATTCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCAATGATCAAAAGTTTCATTAGTGCACCTCATCTCAGAGAGTGAACAATCTCTGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTTCTTCGTGAAATGCAATAGATAGAGAACTCTGGAAGAGCT 979
Db 241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGATAGTCAAGATATGCAAGAAATGACTTTTCGAGGGGACAGAACTCTGGT 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGACAACATTCCTCAAAGCTCGTCTGATTTTGTCTCAGTCCAGGTCCAAATGGC 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGTGATGAACCTGCAGGATGTATTCTTAATGAACCTTTAAAGATCTCTAAA 1159

Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCAGTTGTATATGAGTGTGTTTACGACTTCCAGTAACATTTTCAAGGGATCAGCCGTG 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGATGATGATGAGAGGGTGTTCCTTGGTCCATATGCCACAGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTATCCACGGCCAGGAAC 1339
Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCTCCAGAAACATTTGGTGGTTTGTACTTACAAAGGACCTTCTGTATGATGATTTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTCGAGAGTCCATCCAGCCATGATCAATCCAGTGTTCCTATGAACAATCGCCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACCGATGTAATTAATCAATTTACACAAATTTGCTGTAGACCGAGTGAT 1519
Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 CGAAGATGGAAGATGATGATTTATGTTTATCGGAACAGATGTTGGACCGTCTTTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCATTCCTAAGAGACTTGGTATGATTTAGAGAGGTCTCTCGAAGAAATG 1639
Db 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTTTCGGAACCGACTGCTATTTTCCAGCAATGGAGCTTCCCAATGAAGCAGCAAA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATTGTTTCAACCGCTGGGTTCGCCAGTCCCTTTTACACCGGTGTGATATTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GCGAAGCGGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTCTGGATGGTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCCACTGCAAGAGACGCAAGACGCAAGATATAAGAAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTGTTTCACTTACACATGATGATGATGATGATGATGATGATGATGAT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisGlyHisSerPro 580
QY 1940 GAAGAGAGATCATCTATGCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 1999
Db 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGACGCTGCTTATTTGGCAATTTCCAGAGCGCAATGCAAGAGCAAGAAAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTGGATGATCATATCATCATGAGCAGATGATGATGATGATGATGATGATGATGAT 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAGGATTCAGGCAATTTACTCTCCATGCGGTGGAACATGGGTTCATACAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAACTCGAAGTCTATTTCACAGAGCAATTTTGAAGAACTTCTTCATAAGAT 2239

Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
QY 2240 GATCATCGAGATGCTCTTAAGACCAAGAAATGTCCTAATAGCATGACACCTAGCCAGAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGGTACAGAGACTTTCATCGAGCTCATCAACACACCCCAATCTCAACACGATGATGAG 2359
Db 701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTCAACAAGTTTCGAAAAAGGACCGAAAAACAACGTGCGCAAGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGAACAGTACAAATGGAACCACTTCAACAAAAATAGAAAGGTAGAAAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArg 760
QY 2480 ACCCACTAATTTGAGAGGCGACCCAGGAGTGTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 3
US-09-060-692-54
; Sequence 54, Application US/09060692
; Patent No. 5935865
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,692
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-060-692-54

Alignment Scores: 0 Length: 771
Pred. No.: 4201.00 Matches: 771
Score: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 85.80% Indels: 0
 DB: 2 Gaps: 0
 US-09-774-490-1 (1-2709) x US-09-060-692-54 (1-771)

QY 200 ATGGCTGGTTAACTAGGATGCTGCTCTTTCTGGGAGATATTACTTACAGCAAGACA 259
 DB 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
 QY 260 AACTATCAGAATGGGAAGAACAAATGTGCCAAGCTGAAATATCTCAAGAGAAATGTTG 319
 DB 21 AsnTyrGlnAsnGlyValAsnAsnValProArgLeuLysLeuSerTyrIleGluMetLeu 40
 QY 320 GAATCCAAACAATGTGATCACTTCAATGGCTTGGCCCAAGCTCCAGTTATCATACCTTC 379
 DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
 QY 380 CTTTGGGATGAGGAAGGATAGGCTGTATGTTGGAGCAAGGATCATATTTTCATTC 439
 DB 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
 QY 440 GACCTGGTTAATATCAAGATTTTCAAAAGATTGTCGCCAGTATCTTACACAGAGA 499
 DB 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
 QY 500 GATGAATCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGCTTAATTTTCATCAAGCTA 559
 DB 101 AspGluCysLeuTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
 QY 560 CTTAAGGCATATATCAGACTCATTGTACGCTGTGGAAAGGGGGCTTTTCATCCAAT 619
 DB 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
 QY 620 TGCACCTACATTGAAATTTGGACATCATCTCGAGGACAAATATTTTAACTCGAGAACTCA 679
 DB 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
 QY 680 CATTTTGAAACGGCGTGGGAAGAGTCCATATGACCCCTAAAGCTCTGACAGCATCCCTT 739
 DB 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
 QY 740 TTAATAGATGAGAATTTATCTCTGGAACTGCGAGCTGATTTTATGGCGGAGACTTGTCT 799
 DB 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
 QY 800 ATCTTCCGAATCTTTGGCCACCACCACCAATCAGGACAGAGCATGATTCAGGTGG 859
 DB 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
 QY 860 CTCATGATCCAAAGTTCATTAGTCCCACTCATCTCAGAGAGTGACAAATCCTGAAGAT 919
 DB 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
 QY 920 GACAAAGTATCTTTTCTCGTGAATAATGCAATAGATGAGAACACTCTGGAAAAGCT 979
 DB 241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
 QY 980 ACTCAGCTAGATAGTTCAGATGCAAGATGACTTTGGAGGCGCACAGACTCTGGTG 1039
 DB 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
 QY 1040 AATAAATGGCAACAATTCCTCAAAGCTCGTCTGATTTGCTCAGTGGCCAGGTCCAAATGGC 1099
 DB 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
 QY 1100 ATTGACACTCAATTTTGATGAATCTGAGGATGTATTCCTTAATGAATTTTAAAGATCTTAAA 1159
 DB 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
 QY 1160 AATCCAGTTGTATGGAGTGTTCAGACTTCCAGTACATTTTCAAGGGATCAGCGGTG 1219
 DB 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340

QY 1220 TGTATGTATAGCATCAGTGTGAGAGGGTGTTCCTTGGTCCATATGCCACAGGGAT 1279
 DB 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
 QY 1280 GGACCCAACTATCAATGGTGCTCTTATCAAGAGAGTCCCTATCCAGCGCCAGAACT 1339
 DB 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
 QY 1340 TGTCCAGCAAAAACATTTGGTGGTTTTCAGTCTCAAAAGGACCTTCTCGATGATGTTATA 1399
 DB 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
 QY 1400 ACCTTTCCAGAAGTCATCCAGCCATGTACAAATCCAGTGTTCCTTATGAACAATCGCCA 1459
 DB 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
 QY 1460 ATAGTGATCAAAAACGGATGTAATTTACAAATTTACAAATTTGCTGTAGACGGAGTGTAT 1519
 DB 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
 QY 1520 GCAGAAGATGCACAGTATGATGTTATGTTATCGAACAGATGTTGGACCGTCTCTTAAA 1579
 DB 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
 QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAGAGGTTCTCTCGAAGAAATG 1639
 DB 461 ValValSerIleProLysGluThrTyrAspLeuGluValLeuLeuGluGluMet 480
 QY 1640 ACAGTTTTCGGGAAACCGACTGCTATTTTACAGCAATGGAGCTTTCCACTAAGCAGACAA 1699
 DB 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
 QY 1700 CTATATATTGTTCAACGGCTGGGTTGCCAGCTCCCTTTACACCGGTGTGATTTTAC 1759
 DB 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
 QY 1760 GGGAAAGCGTGTGCTGAGTGTGCTCGCCCGAGACCCCTTACTGTGCTTGGGATGGTCT 1819
 DB 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
 QY 1820 GCATGTTCTCGTATTTTCCCACTGCAAGAGACGCAAGACGACAGATATAGAAT 1879
 DB 541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn 560
 QY 1880 GGACACCCACTGACTCTGTTTCCAGCTTACACCATGATATCACCATGGCCAGCCCT 1939
 DB 561 GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisHisGlyHisSerPro 580
 QY 1940 GAAGAGAGAATCATCTATGTTAGTGTAGAGATAGTAGCACATTTTGGGAATGCGTCCGAG 1999
 DB 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
 QY 2000 TCGCAGAGAGCGTGTCTTATTTGGCAATTCAGAGGCGAAATGAAGACGCAAGAAAGAG 2059
 DB 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
 QY 2060 ATCAGCTGGATGATCATATCATCAGCAGACATCAAGCCCTTCTGCTACGTAGTCTACAA 2119
 DB 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
 QY 2120 CAGAAGGATTCAGGCAATTTACCTCTGCGCATGGTGGAAACATGGGTTCATACAACTCTT 2179
 DB 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
 QY 2180 CTTAAGTAAACCTGGGAAGTCATTGACACAGAGCATTTGGAAGAACTTCTTCAATAAGAT 2239
 DB 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
 QY 2240 GATCATCGAGATGGCTTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCAGCAGAG 2299
 DB 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700

QY 2300 GTCTGTGACAGAGCTTCATGAGCTCATCAACCCCAATCTCAACAGATGATGAG 2359
Db 701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAAGTTTGGAAAAGGACCGGAAACCAACGTCGGCAAGGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTrpLysArgAspA-GlySerGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGAAACAGTAACAATGGAGGACCTTACAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysGlyArgAsnArgArg 760
QY 2480 ACCCAGCAATTTGAGAGGCCACCCAGGAGTGTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 4

US-08-833-391-54
Sequence 54, Application US/08833391
Patent No. 6013781
GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Matthews, David
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,391
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/121,713
FILING DATE: 13-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 771 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-391-54

Alignment Scores:
Pred. No.: 0
Score: 4201.00
Length: 771
Matches: 771
Percent Similarity: 100.00%
Conservative: 0
Best Local Similarity: 100.00%
Mismatch: 0
Query Match: 85.80%
Indels: 0
DB: 3
Gaps: 0

US-09-774-490-1 (1-2709) x US-08-833-391-54 (1-771)

QY 200 ATGGCGTGTAACTAGGATTGTCTGCTTTCTGGGAGTATTACTTACAGCAAGAGA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGATGGGAAGAACAAATGTGCCAAGGCTGAAATATCTCAACAAAGAAATGTTG 319
Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAACATGTATCATTCTCAATGGCTTGGCCAAACAGCTCAGTTATCATCCCTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
QY 380 CTTTGGATGAGGAACGAGTAGGCTGTATCTTCGAGCAAGAGGATCACATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGGCCAGTATCTTACACCAAGAAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATGAATGCAAGTGGCTGGAAAGACATCTCTGAAGAAATGTGCTAATTTTCATCAAGGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCCTGTGACGCTGTGGAACGGGGCTTTTTCATCCAAAT 619
Db 121 LeuLysAlaLysAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProlle 140
QY 620 TGCACCTACATTTGAAATTTGGACATCATCTCTGAGGACAAATATTTTAAAGCTCGAAGACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTGAAAACGGCCGTGGAGAGCTCCATATGACCCCTAAGCTGCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGAGAAATATATCTCTGGAACCTGCACCTGATTTTATGGGCGAGACTTTGCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCCGAACTCTTGGGACCAACCCACCAATCAGGACAGACAGCATGATTCCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProlleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCAATGATCCAAAGTTTCATAGTGGCCACCTCATCTCAGAGAGTGAACATCTCTGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuLysSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGGAGAACACTCTCGGAAAAGCT 979
Db 241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCACGCTAGAATAGGTAGATATGCAAGAATGACTTTTGGAGGCGACAGAAAGTCTGGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGACAACATTCCTCAAGCTCGTCTGATTTGCTCAGTGGCCAGGTCCCAATGGC 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGTATGAATGACCTCAGGATGTATTCCTTAATGAATTTTAAAGATCCTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGTATATGGAGTGTTTTACGACTTCCAGTAAACATTTTCAAGGGGATCAGCGGTG 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGATGTATAGCATGATGATGAGAGGGTGTCTCTTGGTCCATATGCCCCACAGGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCCAACTATCAATGGGTGCCTTATCAAGGAAGAGTCCCTTATCCACGGCCAGGAAC 1339

Db	361	GlyProAsnTyrGlnIrrpValProTyrGlnGlyArgValProTyrProArgProGlyThr	380
Qy	1340	TGTCCCGACAAACATTTGGTGGTTTTGACTCTCAAAAGGACCTTCCTGATGATGTTATA	1399
Db	381	CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle	400
Qy	1400	ACCTTTGCAGAAGTCATCCAGCCATGTACAATCCAGTGTTCCTATGAACAATCGCCCA	1459
Db	401	ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro	420
Qy	1460	ATAGTGATCAAAACGGATGTAAATATCAATTTACAAATTTGCTAGACCGAGTGGAT	1519
Db	421	IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp	440
Qy	1520	GCAGAAGATGCACAGTATGATGTTATGTTATTCGGAACAGATGTGGACCGTTCCTTAAA	1579
Db	441	AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys	460
Qy	1580	GTAGTTTCAATTCCTAAGGAGACTTGATATGATTTAGAAGAGGTTCGTGCGAAGAAATG	1639
Db	461	ValValSerIleProLysGluThrTyrTyrAspLeuGluValLeuLeuGluGluMet	480
Qy	1640	ACAGTTTTTCGGGAACCGACTGCTATTTCAGCAATGAGCTTTCACCTAAGCAGCAACAA	1699
Db	481	ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln	500
Qy	1700	CTATATATTGGTTCAAACGGCTGGGGTTGCCAGCTCCCTTTTACACCGGTGTGATATTTC	1759
Db	501	LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr	520
Qy	1760	GGGAAACGCTGTCTGAGTGTGCTCCGCCGAGACCTTACTGTGCTGGGATGGTTCT	1819
Db	521	GlyLysAlaCysAlaGluCysLysLeuAlaArgAspProTyrCysAlaIleTyrAspGlySer	540
Qy	1820	GCATGTTCTCGCTATTTTCCCACTGCGAAAGAGCGCAAGACAAGAATATAGAANAAT	1879
Db	541	AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn	560
Qy	1880	GGAGACCACTGACTCACTGTTTACAGCTTACACCATGATATACCATGCGCCACAGCCCT	1939
Db	561	GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro	580
Qy	1940	GAAGAGAGAATCATCTATGTTGTAGAGAAATAGTAGCAATTTTTGGAAATGCAGTCCGAAG	1999
Db	581	GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys	600
Qy	2000	TCGCAGAGAGCGCTGCTTATTGGCAATTCAGAGCGCAATGAGAGCGCAAAAGAGAGAG	2059
Db	601	SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu	620
Qy	2060	ATCAGATGGATGATCATATCATCAGACAGATCAAGCGCTTCCTGCTACCTAGTCTACAA	2119
Db	621	IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln	640
Qy	2120	CAGAAGGATTTCAGCAATTAACCTCTGCCATCGGTGGAAACATGGGTTTCATACAACTCTT	2179
Db	641	GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu	660
Qy	2180	CTTAAGGTAAACCTGGGAAGTCATTGACACAGCATTTTGAAGAACTTCTTCATAAAGAT	2239
Db	661	LeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLysAsp	680
Qy	2240	GATGATCGAGATGGCTTAAGACCAAGAATGTCCAATAGCATGACACCTACGCCAGAG	2299
Db	681	AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys	700
Qy	2300	GTCTGGTACAGACTTTCATGCACTCATCAACCAACCCCAATCTCAACACGATGGATGAG	2359
Db	701	ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu	720
Qy	2360	TTCTGTGAACAAGTTTGGAAAGGACCGGAAACCAACGTCGGCAAGGCCGAGACATACC	2419

```

Db      721  PheCy6gluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY      2420  CCAGGGAACAGTAAACAAATCGAAGCACTTACAGAAATAAGAAAGGTAGAAACAGGAGG 2479
          |||||
Db      741  ProGlyAsnSerAsnLysTrpLysHisLeuGlnGlnLysLysGlyArgAsnArgArg 760
          |||||
QY      2480  ACCACAGAAATTGAGAGGGCACCCAGGAGTGTC 2512
          |||||
Db      761  ThrHisGluPheGluArgAlaProArgSerVal 771
          |||||

RESULT 5
US-09-060-610-54
; Sequence 54, Application US/09060610
; Patent No. 6344544
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,610
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/835,268
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
;
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-060-610-54

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 4201.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 85.80% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-060-610-54 (1-771)

QY      200  ATGGGCTGGTTAACTAGGATTCTGTCTTTCTGGGGAGTATTACTTACAGCAAGCA 259
          |||||
Db      1  MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
          |||||
QY      260  AACTATCAGAAATGGGAAGAACAAATGTGCCAAGCGCTCAAATTATCTTACAAAGAAATGTTG 319
          |||||

```

21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
320 GAATCCACAAGTGTATCACTTTCAATGGCTGGCCACAGCTCCAGTATATACACCTTC 379
41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
380 CTTTGGTATGAGCAACGAGTAGGTGTATGTTGGAGCAAGGATCACATATTTTCATTC 439
61 LeuLeuAspGluLysArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
440 GACCTGGTTAAATATCAAGATTTTCAAAAGATGTGTGGCCAGTATCTTACACCAAGA 499
81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
500 GATGAATGCAAGTGGCTGGAAGAAGACATCCTGAAGATGTCTAAATTTTCATCAAGTA 559
101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
560 CTTAAGGCATATAATCAGACTCACTGTACGCTGTGGAAACGGGGCTTTTTCATCAATT 619
121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
620 TCACCTACATTAATCGACATCATCTGAGGACAATATTTTAAAGCTGGAGAACTCA 679
141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
680 CATTGTGAACGGCGTGGAGAGTCCATATGACCTTAAGCTGCTGACAGCATCCCTT 739
161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
740 TTAATAGATGGAGATTAATATCTGGAACCTGAGCTGTATTTATGGCGAGACTTTGCT 799
181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
800 ATCTCCGCAACTCTGGGCACCAACCAATCAGACAGACAGACAGATTCACAGGTGG 859
201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
860 CTCATATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGACAACTCTCAAGT 919
221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
920 GACAAAGTATATCTTTCTCCGTGAAATGCAATAGATGGAGAACACTCTGGAAGCT 979
241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyHisSerGlyLysAla 260
980 ACTCAGCTAGATAGTGCAGATATGCAAGAATGACTTTGGAGGCGCAGAAAGTCTGGT 1039
261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
1040 AATAATGACACATCTCCAAAGCTGCTGATTTGCTCAGTCCAGGTCCCAATGGC 1099
281 AsnLysTrpThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
1100 ATTGACACTCATTTTGATCAACTGCAAGATGTATTCCTAATGAACTTTAAAGATCCTAAA 1159
301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
1160 AATCCAGTTGTATGAGTGTATGACTTCAGACTTCAGTAAACATTTTCAAGGATCACCGGT 1219
321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
1220 TGTATGTATAGATGATGATGAGAGGTGTTCTGTGTCATATGCCACACAGGAT 1279
341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
1280 GGACCCCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTATCCAGCGCCAGGAACT 1339
361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
1340 TGTCCAGCAAAACATTTGGTGGTTTGTACTACAAAGGACCTTCTGTATGATGTTATA 1399
381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400

1400 ACCTTTGCAAGAAAGTCATCCAGCCATGTACAATCCAGTGTTCCTTATGAAACAATGCCCA 1459
401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
1460 ATAGTGTCAAAAACGAGTAAATATCAATTTTACAAAATTTGCTGTAGACCGAGTGGAT 1519
421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
1520 GCAGAAGATGCAGATGATGATTTATGTTTATCCGACACAGATGTGGACCGTCTCTTAAA 1579
441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTTAGAAGAGGTTCCTCTCGAAGAAATG 1639
461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluGluMet 480
1640 ACAGTTTTTCGGGAACCGACTGTCTATTTCAGCAATGGAGCTTTTCCATTAAGCAGACAA 1699
481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
1700 CTATATATTGGTTCAACGGCTGGGTTGCCAGCTCCCTTTTACACCGGTGTGATATTTAC 1759
501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
1760 GCGAAAGCGTGTGCTGAGTGTGCTCGCCGACAGACCTTACTGTGCTTGGATGGTTCT 1819
521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
1820 GCATGTTCTCGCTATTTCACACTGCAAGAGAGACGCAAGACGCAAGATATAGAAT 1879
541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn 560
1880 GGACACCCACTGACTCTACTGTTACACTTACACCATGATATCACCATGGCCACAGCCCT 1939
561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580
1940 GAACAGAGAATCATCTATGTTGAGATAAGTAGACACATTTTGGAAATGACGATCCCGAAG 1999
581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
2000 TCGCAGAGAGCGTGTGCTTATGTCATTCAGAGGCGAATGAAGACGCAAGAAAGAGAG 2059
601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluGluArgLysGluGlu 620
2060 ATCAGAGTGGATGATCATATCATCAGACAGATCAAGSCCTTCTGCTACGTAGTCTACAA 2119
621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
2120 CAGAAGGATTCAGGCAATTTACTCTGCTCATCGGTGGAACATGGGTTTCATACAACTCTT 2179
641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
2180 CTTAAGGTAAACCTCGAAGTCAATTCACAGAGCAATTTGGAAGAACTTCTTCATAAAGAT 2239
661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
2240 GATCATGAGATGCTCTTAGACCAAAAGATGTCCAATAGCATGACACCTAGCCAGAGAG 2299
681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
2300 GTCTGGTACAGACTTTCATGCTCATCAACCAACCCCAATCTCAACACCATGGATGAG 2359
701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
2360 TTCTGTGAACAAGTTTGGAAAAGGACCGAAGAAACACAGTCGCGCAAGCCAGGACATACC 2419
721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgArgGlnArgProGlyHisThr 740
2420 CCAGGACAGTAAACAATGGAACACTTACAAGAAATAGAAGGTAGAGAAACAGAGAGG 2479
741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760

QY 2480 ACCCAGCAATTTGAGAGGCGCCAGAGTGTCTC 2512
 Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 6
 PCT-US94-10151A-54
 ; Sequence 54, Application PC/TUS9410151A
 ; GENERAL INFORMATION:
 ; APPLICANT: The Regents of the University of California
 ; TITLE OF INVENTION: The Semaphorin Gene Family
 ; NUMBER OF SEQUENCES: 66
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: FLEHR HOEBACH TEST ALBRITTON & HERBERT
 ; STREET: 4 Embarcadero Center, Suite 3400
 ; CITY: San Francisco
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94111-4187
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US94/10151A
 ; FILING DATE: 13-SEP-1994
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Osman, Richard A.
 ; REGISTRATION NUMBER: 36,627
 ; REFERENCE/DOCKET NUMBER: FP-58750-PC/RAO
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 781-1989
 ; TELEFAX: (415) 398-3249
 ; TELEX: 910 277299 FHT UR
 ; INFORMATION FOR SEQ ID NO: 54:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 771 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; PCT-US94-10151A-54

Alignment Scores:
 Pred. No.: 0 Length: 771
 Score: 4201.00 Matches: 771
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 85.80% Indels: 0
 DB: 5 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US94-10151A-54 (1-771)

QY 200 ATGGCTGGTAACTAGAGATGTCTGCTTTCTGGGAGTATTACTTACAGCAAGACA 259
 Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20

QY 260 AACTATCAGAAATGGGAAGAACAAATGTCCAGGCTGAAATTTATCTCAAGAAAGAAATGTTG 319
 Db 21 AsnTyrGlnAsnGlyLysAsnValProArgLeuLysLeuSerTyrLysGluWetLeu 40

QY 320 GAATCCAAATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCATACCTTC 379
 Db 41 GluSerAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60

QY 380 CTTTGGATGAGGACCGAGTAGCTGTATGTTGGAGCAAGAGTACATATTTTCATTC 439
 Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80

QY 440 CACCTGGTTAATATCAAGGATTTTCAAAGATTTGTGGCCAGTATCTTACACCAAGAGA 499
 Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100

QY 500 GATGAATGCAAGTGGGCTGGAAAAAGACATCTCTGAAAGAAATGTGCTAATTTTCATCAAGGTA 559
 Db 101 AspGluCysIysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120

QY 560 CTTAAGGCATATAATCAGACTCAGTGTGACCCCTGTGGAACGGGGCTTTTCATCCAATT 619
 Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140

QY 620 TGCACCTACATTTGAAATTTGGACATCATCTCTGAGGACAAATATTTTAACTCTGGAGACTCA 679
 Db 141 CysThrTyrIleGluIleGlyHisHisProGluAsnIlePheLysLeuGluAsnSer 160

QY 680 CATTTTGAACGGCGCTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTT 739
 Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180

QY 740 TTAATAGATGGAGAAATATATCTCTGGAACTCCAGCTCAGCTGATTTTATGGGCGGAGACTTCT 799
 Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200

QY 800 ATCTTCCGAACTCTTGGGCACCCACCACCAATCAGGACAGAGCAGCATGATTTCCAGGTGG 859
 Db 201 IlePheArgThrLeuGlyHisHisHisProIleArgThrGluGlnHisAspSerArgTrp 220

QY 860 CTCATATGATCCAAAGTTTCATTTAGTGGCCACCTCATCTCAGAGAGTGACAAATCTCGAAGAT 919
 Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240

QY 920 GACAAAGTATACTTTTCTCCGTAAGTCAAGAAATGCAATAGATGAGACACTCTCGAAAGCT 979
 Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260

QY 980 ACTCAGCTGAGATAGTGTGATGATGCAAGAAATGCACTTTTGGAGGCGACAGAAAGTCTGTGT 1039
 Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280

QY 1040 AATAAATGGACAAATCTCCAAAGCTCGTCTGATTTGCTCAGTGCCAGGTCCCAATGCGC 1099
 Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300

QY 1100 ATTGACACTCAATTTTGTATGAGTGCAGAGTATCTCTATGAACTTTTAAAGATCTTAAA 1159
 Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320

QY 1160 AATCCAGTTGTATGAGTGTGTTTACGACTTCCAGTAAACATTTTCAAGGGATCAGCCGTG 1219
 Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340

QY 1220 TGTATGTATGATGATGATGTGAGAGGGTGTCTCTTGGTCCATATGCCCAAGGAT 1279
 Db 341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360

QY 1280 GGACCCCACTCATCAATGGTGCCTTATCAAGGAGAGTCCCTATCCAGGCCAGGAACT 1339
 Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380

QY 1340 TGTCCCAAGCAAAACATTTGGTGGTTTGTACTCTCAAGAGACCTTCTCTGATGATGTATA 1399
 Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400

QY 1400 ACCTTTTCAAGAAAGTCAATCCAGCCCATGACAATCCAGTGTCTCTATGAACAATCGCCCA 1459
 Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420

QY 1460 ATAGTGATCAAAACCGAGTGAATTTATCAATTTTACAAATTTGCTGAGCCGAGTGGAT 1519
 Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440

QY 1520 GCAGAAGATGCAGATGATGATGTTATGTTTATCGGAACAGATGTTGGGCGCGTTCCTTAAA 1579
 Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460

QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGTGTATGATTTTAGAAGAGGTTCTGCTCGAAGAAATG 1639

```
|||||ValValSerIleProLysGluThrTrpTyAspLeuGluValLeuLeuGluMet 480
QY 1640 ACAGTTTTCGGGAACCGACTGCTATTTCAGCAATGGAGCTTTCACATAGCAGCAACA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATTGGTTCAACGCTCGGGTTCGCCAGCTCCCTTACACCGGTGTATATTAC 1759
Db 501 LeuTyriIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTy 520
QY 1760 GCGAAAGCGTGCTGCTGAGTGTTCCTCGCCGAGACCCCTTACTGCTGGGATGTTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyCysAlaTrpAspGlySer 540
QY 1820 GATGTTCTCGCTATTTTCCCACTGCAAGAGACGACAGCAGCAGATATAAGAAAT 1879
Db 541 AlaCysSerArgTyPheProThrAlaLysArgArgThrArgGlnAspIleArgAsn 560
QY 1880 GGAGACCACTGACTGCTTCACTTACACTTACACCTGATTAATCACCATGCCACAGCCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisIleGlyHisSerPro 580
QY 1940 GAAGAGAGAATCATCTATGCTGTAGAGATAGTACACATTTTTCGAATGCAGTCCGAAG 1999
Db 581 GluGluArgIleIleTyGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TGCAGAGAGCGCTGCTATTGGCAATTCAGAGCGCAATGAAGAGCAAAAGAGAG 2059
Db 601 SerGlnArgAlaLeuValTyTrpGlnPheGlnArgArgAsnGluGluArgLysGlu 620
QY 2060 ATCAGAGTCGATCATATCATCAGCAGACATCAGGCTTCTGCTAGTGTCTACAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGGATTTCAGGCAATTACCTCTGCCATGCGGTGGAACATGGGTTCATACAAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAACTGGAAGTCAATTCACAGAGCAATTTGGAAAGAACTTCTTCATAAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuLeuHisLysAsp 680
QY 2240 GATGATGGAGATGGCTCTAAGACCAAGAAATGTCATAGCATGACACCTAGCCAGAG 2299
Db 681 AspAspGlyAspLysSerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGTACAGAGACTTCATGAGCTCATCAACCCCAATCTCAACACGATGGATGAG 2359
Db 701 ValTrpTyArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAAGTTTGGAAAGGGCCGAAACACGTCGCAAGGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGAACAGTAACAAATGGAGCACTTACAGAAATAAAGAGGTAGAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysLysArgAsnArgArg 760
QY 2480 ACCCAGAAATTTGAGAGGCGCCAGGAGTGTCT 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771
```

RESULT 7

```
US-08-556-422A-3
; Sequence 3, Application US/08556422A
; Patent No. 6576754
; GENERAL INFORMATION:
; APPLICANT: HALL, Kathryn T.
; APPLICANT: FREEMAN, Gordon J.
; APPLICANT: SCHULTZE, Joachim L.
; APPLICANT: BOUSSIOTIS, Vassiliki
; APPLICANT: NADLER, Lee M.
```

```
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES
; FILE REFERENCE: DFN-005CPA2
; CURRENT APPLICATION NUMBER: US/08/556.422A
; CURRENT FILING DATE: 1995-11-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-556-422A-3
Alignment Scores:
Pred. No.: 0 Length: 655
Score: 3568.00 Matches: 655
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 72.88% Indels: 0
Db: 4 Gaps: 0
```

US-09-774-490-1 (1-2709) x US-08-556-422A-3 (1-655)

```
QY 200 ATGGCGGTAACTAGAGATTGCTGCTTTCTGGGAGTATTACTTACAGCAGAGACA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGATGGGAACAATGTCCAAGGCTGAAATATCTCTACAAAGAAATGTTG 319
Db 21 AsnTyriGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyLysGluMetLeu 40
QY 320 GAATCCAACTATGATGATCATTTCATCAATGGCTTGGCCAAACAGCTCCAGTTATCATCCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyHisThrPhe 60
QY 380 CTTTGTGATAGGAACGAGTAGCTGTATGTTGGAGCAAGAGTACATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACTGTGTTAATATCAAGGATTTTCAAAGATTCTGTGCCAGTATCTTACACCAAGAAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyTrpArgArg 100
QY 500 GATCAATGCAAGTGGGTGGAAAGACATCTCTGAAAGAAATGTCTTAATTTTCATCAAGTGA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATATACAGACTCATTGTACCCCTGTGGAACGGGGCTTTTCATCCCAATT 619
Db 121 LeuLysAlaTyAsnGlnThrHisLeuTyAlaCysGlyThrGlyAlaPheHisPheIle 140
QY 620 TGCACCTACATTGAAATTTGGACATCATCTCGAGGACCAATATTTTAACTCGAGAACTCA 679
Db 141 CysThrTyriIleGluIleGlyHisIleProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAAAACGGCCGTGGGAAGAGTCCATATGACCCCTTAAGCTGCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyArgProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGAGAAATATATCTCTGGAACCTGACCTGATTTTATGGGGCAGACTTTCCT 799
Db 181 LeuIleAspGlyGluLeuTySerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCCGAACTCTGGGACCCACCCATCAGGACAGCAGAGTATTCAGAGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisIleProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCAATGATCCAAAGTTTCATTAGTGGCCACCTCATCTCAGAGAGTGAACAATCTCGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTTCTCCGTGAAATGCAATAGTAGAGAACACTCTGGAAAGCT 979
Db 241 AspLysValTyPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
```

QY 980 ACTCAGCTAGATAGTTCAGATATGCAAGATGACTTTTGGAGGCGCAGAGAAGTCTGGTG 1039
Db ThrHisAlaArgIleGlyGlnIleCysLysAsnAepPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGCAACAATCTCTCAAGCTCGTCTGATTGCTCAGTGCAGGCTCCAAATGGC 1039
Db AsnLysIleThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCAATTTTGATGAAGTGCAGGATGATTTCTTAATGAACCTTTAAAGATCTTAAA 1159
Db IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCAGTTGTATGAGTGTTCAGACTCCAGTAAACATTTCAAGGATCAGCGGTG 1219
Db AsnProValValTyGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATACATGAGTGTGAGAGGGTGTTCCTTGTCTCATATGCCAGGAGAT 1279
Db CysMetTySerMetSerAspValArgArgValPheLeuGlyProTyAlaHisArgAsp 360
QY 1280 GGACCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCACGGCCAGGAAT 1339
Db GlyProAsnTyGlnTrpValProTyGlnGlyArgValProTyProArgProGlyThr 380
QY 1340 TGTCCAGCAAAACATTTGGTGGTTTGAATCTCAAGAGACCTTCTCGATGATGTTATA 1399
Db CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAGAACTCATCCAGCATGTACAAATCCAGTGTTCCTATGAACAATCGGCCA 1459
Db ThrPheAlaArgSerHisProAlaMetTyArgProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACCGGATGTAATATCAATTTTACAAATTTGTCTAGACCGAGTGGAT 1519
Db IleValIleLysThrAspValAsnTyGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAGATGACAGTATGATGTTATGTTATCGGAACAGATGTTGGACCGTCTTTAAA 1579
Db AlaGluAspGlyGlnTyArgValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATCTCAGGACTGTGTATGATTAATAGAGAGTTCGTCTGGAGAAGATG 1639
Db ValValSerIleProLysGluThrTrpTyArgLeuGluValLeuLeuGluMet 480
QY 1640 ACAGTTTTTCCGGAAACCGACTGTATTTACAAATTTGTCTAGACCGAGTGGAT 1699
Db ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
QY 1700 CTATATATGTTTCAAGCTGGGTGGTCCAGCTCCCTTTACACCGGTGTGATATTATC 1759
Db LeuTyIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTy 520
QY 1760 GGGAAAGCTGTCTGAGTGTTCCTCGCCGAGACCTTACTGTCTGGATGGTCT 1819
Db GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCCCTATTTTCCCACTGCAAGAGAGCGCAAGACGACAGATATAGAAT 1879
Db AlaCysSerArgTyPheProThrAlaLysArgThrArgArgGlnAspIleArgAsn 560
QY 1880 GGAGACCACTGACTCTGTTGAGCTTACACCATGATATCAACCATGGCCAGCCCT 1939
Db GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisHisGlyHisSerPro 580
QY 1940 GAAGAGAGATCATCTATGTTGAGTATAGATAGTACACATTTTGGATGCGATCCGAG 1999
Db GluGluArgIleIleTyGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGAGCGCTGCTTATTTGGCAATTCAGAGGCGAAATGAAGACGCAAGAGAG 2059
Db SerGlnArgAlaLeuValTyTrpGlnPheGlnArgAsnGluGluArgLysGluGlu 620

QY 2060 ATCAGAGTGGATGATCATATCATCAGGACAGATCAAGCCCTTCTGCTAGTACTACAA 2119
Db 621 IleArgValAspAspHisIleIleLeuArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAGAGATTCAGCAATTAACCTCTGCATGCGGTGGAGACATGGG 2164
Db 641 GlnLysAspSerGlyAsenTyLeuCysHisAlaValGluHisGly 655
RESULT 8
US-08-136-922-2
; Sequence 2, Application US/08136922
; Patent No. 5416197
; GENERAL INFORMATION:
; APPLICANT: Raper, Jonathan A.
; APPLICANT: Luo, Yuling
; TITLE OF INVENTION: Compositions Which Regulate Neural
; TITLE OF INVENTION: Regeneration and Methods of Making and Using the Same
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: NO. 541619718
; STREET: One Liberty Place
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/136,922
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: UPN-1428
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-136-922-2
Alignment Scores:
Pred. No.: 1,01e-246 Length: 477
Score: 2562.00 Matches: 472
Percent Similarity: 99.37% Conservative: 0
Best Local Similarity: 99.37% Mismatches: 3
Query Match: 52.33% Indels: 0
DB: 1 Gaps: 0
US-09-774-490-1 (1-2709) x US-08-136-922-2 (1-477)
QY 899 GAGAGTGACAATCCTGAAGATGACAAAGTATATCTTTTCTTCGTGAAATGCAATAGAT 958
Db 2 GluHisAspAsnProGluAspAspLysValTyPhePheArgGluAsnAlaIleAsp 21
QY 959 GGAGACACTCTGAAAAGCTACTCAGCTAGATAGTTCAGATATCCAGAAATGACTTT 1018
Db 22 GlyGluHisSerGlyLysAlaThrHisAlaArgIleGlyGlnIleCysLysAsnAspPhe 41
QY 1019 GGAGGGCACAAGATCTGGTGAATTAATGACAACATTCCTCAAGCTCGTCTGATTGTC 1078
Db 42 GlyGlyHisArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeuIleCys 61
QY 1079 TCAGTGCCAGGTCCTCAAAATGCGATTCATCTATTTTATGAACTGCAGGATGATTCTCTA 1138

```
Db 62 SerValProGlyProAsnGlyIleAspThrHisPheAspGluLeuGlnAspValPheLeu 81
QY 1139 ATGAACCTTTAAGATCCTAAATAATCCAGTTGTATATGAGAGTGTGTACAGCTTCGATTAAC 1198
Db 82 MetAsnPheLysAspProLysAsnProValValTyrGlyValPheThrSerSerAsn 101
QY 1199 ATTTTCAAGGATCAGCGTGTATGTATATAGCATGATGTAGTGTAGTGTAGTGTGTGTTCCTT 1258
Db 102 IlePheLysGlySerAlaValCysMetTyrSerMetSerAspValArgValPheLeu 121
QY 1259 GGTCCATATGCCACAGGAGTGGACCCCAACTATCAATGGGTGGCTTATCAAGGAAGAGTC 1318
Db 122 GlyProTyrAlaHisArgAspGlyProAsnTyrGlnTrpValProTyrGlnGlyArgVal 141
QY 1319 CCCTATCACGGCCAGGAACTGTGCCAGCAAAACATTTGGTGGTTTGGATCTACAAAG 1378
Db 142 ProTyrProArgProGlyThrCysProSerLysThrPheGlyGlyPheAspSerThrLys 161
QY 1379 GACCTTCCTGATGATGTTATACCTTTCGAAGAAGTCATCCAGCCATCAAGTCAATCCAGTG 1438
Db 162 AspLeuProAspAspValIleThrPheAlaArgSerHisProAlaMetTyrAsnProVal 181
QY 1439 TTTCTCATCAACAATCGCCCAATAGTATCAAAACGGATGTAATAATTATCAATTTACAAA 1498
Db 182 PheProMetAsnAsnArgProIleValIleLysThrAspValAsnTyrGlnPheThrGln 201
QY 1499 ATTGCTGTAGCCAGTGTGATCGAAGATGACAGATGATGATGATGATGATGATGATGATG 1558
Db 202 IleValValAspArgValAspAlaGluAspGlyGlnTyrAspValMetPheIleGlyThr 221
QY 1559 GATGTTGGGACCGTTCCTAAAGTACTTCAATCTCAAGGAGACTTGTGATGATTTAGAA 1618
Db 222 AspValGlyThrValLeuLysValValSerIleProLysGluThrTrpIlyAspLeuGlu 241
QY 1619 GAGGTTCTGCTGGAAGAAATCACAGTCTTTTCGGGAACCGACTGCTATTTACGCAATGAG 1678
Db 242 GluValLeuLeuGluMetThrValPheArgGluProThrAlaIleSerAlaMetGlu 261
QY 1679 CTTTCCACTAGCAGCAACAATATATATGTTTCAACGGCTGGGGTGGCCAGCTCCCT 1738
Db 262 LeuSerThrLysGlnGlnLeuTyrIleGlySerThrAlaGlyValAlaGlnLeuPro 281
QY 1739 TTACACCGTGTGATATTTACGGGAAGCGTGTGCTGATGTGCTCGCCCGAGACCT 1798
Db 282 LeuHisArgCysAspIleTyrGlyLysAlaCysAlaGluCysLysLeuAlaArgAspPro 301
QY 1799 TACTGTCTTGGGATGGTTCTGCTGCTTCTGCTATTTTCCACTGCAAGAGAGCGACA 1858
Db 302 TyrCysAlaTrpAspGlySerAlaCysSerArgTyrPheProThrAlaLysArgArgThr 321
QY 1859 AGACGACAGATATAAGAAATGGAGACCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1918
Db 322 ArgArgGlnAspIleArgAsnGlyAspProLeuThrHisCysSerAspLeuHisAsp 341
QY 1919 AATCACCATGCCACAGCCCTGAACAGAGAAATCATCTATGTTGTAGAGAAATAGTAGACA 1978
Db 342 AsnHisLysGlyHisSerProGlnGluArgIleIleTyrGlyValGluAsnSerSerThr 361
QY 1979 TTTTGGAAATGCAGTCCGAAGTCGACAGAGCGCTGCTGCTATTTGGCAATTCGAGGCGA 2038
Db 362 PheLeuGluCysSerProLysSerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArg 381
QY 2039 AATGAAGCGGAAAGAGATCATGAGTGTGATGATGATGATGATGATGATGATGATGATGAT 2098
Db 382 AsnGluGluArgLysGluGluIleArgValAspAspHisIleIleArgThrAspGlnGly 401
QY 2099 CTTTCTGCTAGCTAGCTCAACAGAGGATTCAGGCAATTAACCTCTGCTGCTGCTGCTGCT 2158
Db 402 LeuLeuLeuArgSerLeuGlnGlnLysAspSerGlyAsnTyrLeuCysHisAlaValGlu 421
QY 2159 CATGGGTTCATCAAAACTTCTTAAAGGTAAACCTTGAAGTCAATTCACACAGAGCATTTG 2218
```

```
Db 422 HisGlyPheIleGlnThrLeuLeuLysValThrLeuGluValIleAspAsnGluHisLeu 441
QY 2219 GAAGAACTTCTTCATAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2278
Db 442 GluGluLeuLeuHisLysAspAspGlyHisGlySerLysThrLysGluMetSerAsn 461
QY 2279 AGCATGACACCTACCCAGAAGTCTGCTACAGAGACTTCATGCAG 2323
Db 462 SerMetThrProSerGlnLysValTrpTyrArgAspPheMetGln 476

RESULT 9
US-09-308-179B-1
; Sequence 1, Application US/09308179B
; Patent No. 6436669
; GENERAL INFORMATION:
; APPLICANT: INAGAKI, Shinobu
; APPLICANT: FURUYAMA, Tatsuo
; TITLE OF INVENTION: NOVEL SEMAPHORIN GENES (I)
; FILE REFERENCE: 0020-4562P
; CURRENT APPLICATION NUMBER: US/09/308.179B
; CURRENT FILING DATE: 1999-05-14
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: PCT/JF97/04111
; PRIOR FILING DATE: 1996-11-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-308-179B-1

Alignment Scores:
Pred. No.: 1,54e-193 Length: 775
Score: 2030.00 Matches: 368
Percent Similarity: 67.75% Conservative: 155
Best Local Similarity: 47.67% Mismatches: 231
Query Match: 41.46% Indels: 18
DB: 4 Gaps: 8

US-09-774-490-1 (1-2709) x US-09-308-179B-1 (1-775)
QY 218 ATTGCTGTCTTTCTCTGGGAGTATTACTTACAGCAAGAGCAAACTATCAAGATGGGAAG 277
Db 8 LeuThrLeuLeuLeuTrpGlyHisLeuLeuGluLeuThrProGlyHisSerAlaAsn 27
QY 278 AACAAATGCGCAAGGCTGAAATATCTTACAAAGAAATGTTGGAATCCAAATGTGATC 337
Db 28 ProSerTyrProArgLeuArgLeuSerHisLysGluLeuLeuLeuLeuLeuLeuLeuLeu 47
QY 338 ACTTTCATGCTGGCCCAACAGCTCCAGTATCATACCTTCCTCTTGGATGAGGAACGG 397
Db 48 IlePheGlnSerProLeuGlyPheLeuAspLeuHisThrMetLeuLeuAspGluTyrGln 67
QY 398 AGTAGGCTGTATGTGGCAAGAGATCACATATTTTCATTCGACCTGGTTAAATATCAAG 457
Db 68 GluArgLeuPheValGlyArgAspLeuValTyrSerLeuLeuLeuLeuLeuLeuLeuLeu 87
QY 458 GAT---TTTCAAAAGATGTTGGCCAGTATCTTACACCAAGAGAGATGAATGCAAGTGG 514
Db 88 AspGlyTyrArgGluIleTyrTrpProSerThrAlaValLysValGluGluCysIleMet 107
QY 515 GCTCGAAAGACATCCGAAAGATGCTAATTTTCATCAAGGACTTAAGGCATATTAAT 574
Db 108 LysGlyLysAsp---AlaAsnGluCysAlaAsnTyrIleArgValLeuHisThrAsn 126
QY 575 CAGACTCACTGTACGCTGTGGAAACGGGGCTTTTCATCCAAATTCACCTACATTCGAA 634
Db 127 ArgThrHisLeuLeuThrCysAlaThrGlyAlaPheAspProHisCysAlaPheIleArg 146
QY 635 ATTGCAATCATCTCTAGGAGCAATATTTTAAAGCTGGAGAACTCACATTTTGAAGAACGGC 694
```

Db 147 ValGlyHisHisSerGluGluProLeuPhePheHisLeuGluSerHisArgSerGluArgGly 166
Qy 695 CQTGGAAAGAGTCCATATGACACCTTAAGCTGCTGACAGCATCCCTTTTAAATAGATGAGAA 754
Db 167 ArgGlyArgCysProPheAspProAsnSerSerPheValSerThrLeuValGlyAsnGlu 186
Qy 755 TTATACTCTGGAACTGACAGCTGATTTTATGGGCGAGACTTGTCTATCTTCGAACTCTT 814
Db 187 LeuPheAlaGlyLeuTyrSerAspTyrTrpGlyArgAspSerAlaIlePheArgSerMet 206
Qy 815 GGGCACCACCCACCAATCAGGACAGCAGCATGATCCAGGTGCTCAATGATCCAAAG 874
Db 207 GlyLysLeuGlyHisIleArgThrGluHisAspAspGluArgLeuLeuLysGluProLys 226
Qy 875 TTCATTAGTGCCACCTCATCTCAGAGATGACAACTCTGGAAGATGACAAAGTATACTTT 934
Db 227 PheValGlySerTyrMetIleProAspAsnGluAspArgAspAsnLysMetTyrPhe 246
Qy 935 TTCTCCGTGAATGCAATAGATGAGAACACCTCTGGAAGAGCTACTCAGCTAGATA 994
Db 247 PhePheThrGluLysAlaLeuGluAlaGluAsnAsnAlaHisThrIleTyrThrArgVal 266
Qy 995 GGTCAATATGCAAGAATGACTTTTCGAGGCGACAGAAGCTCTGCTGAATAAATGACACAA 1054
Db 267 GlyArgLeuCysValAsnAspMetGlyGlyGlnArgIleLeuValAsnLysTrpSerThr 286
Qy 1055 TTCTCAAAGCTCGTCTGATTTGCTCAGTCCAGGTCCAAATGCGATTGACACATCTTT 1114
Db 287 PheLeuLysAlaArgLeuValCysSerValProGlyMetAsnGlyIleAspThrTyrPhe 306
Qy 1115 GATGAAGTCAGGATGATTTCTTAATGAATTTAAAGATCTCTAAATATCCAGTTGTATAT 1174
Db 307 AspGluLeuGluAspValPheLeuLeuProThrArgAspProLysAsnProValIlePhe 326
Qy 1175 GGAGTGTTCAGACTTCCAGTAACAATTTTCAAGGATCAGCGCTGTGATGTATAGATG 1234
Db 327 GlyLeuPheAsnThrThrSerAsnIlePheArgGlyHisAlaValCysValTyrHisMet 346
Qy 1235 AGTATGTGAGAGGTTGCTCTTGGTCCATATGCCCCAGGATGGACCCCACTATCAA 1294
Db 347 SerSerIleArgGluAlaPheAsnGlyProTyrAlaHisLysGluGlyProGluTyrHis 366
Qy 1295 TGGTGCCTTTATCAAGAAAGATGCCCTATCCAGGCCAGAACTTGTCCCGCAAAACA 1354
Db 367 TrpSerLeuTyrGluGlyLysValProTyrProArgProGlySerCysAlaSerLysVal 386
Qy 1355 TTGTGTGT---TTGACTCTACAAAGACCTTCTGATGATGTATTAACCTTTGCCAGA 1411
Db 387 AsnGlyGlyLysTyrGlyThrThrLysAspTyrProAspAspAlaIleArgPheAlaArg 406
Qy 1412 AGTCATCCAGCCATGTATCAATCCAGTGTTCCTATGAACAATCGCCCAATAGTATCAA 1471
Db 407 MethHisProLeuMetTyrGlnProIleLysProValHisLysLysProIleLeuValLys 426
Qy 1472 ACGGATGTAATTTCAATTTACAAATTTGCTGTAGACCGAGTGGATGCGAGAGATGGA 1531
Db 427 ThrAspGlyLysTyrAsnLeuArgGlnLeuAlaValAspArgValGluAlaGluAspGly 446
Qy 1532 CAGTATGATGTATTTTATCGGAACAGACTTGTGGACCGTCTTAAAGTAGTTCAATT 1591
Db 447 GlnTyrAspValLeuPheIleGlyThrAspThrGlyIleValLeuLysValIleThrIle 466
Qy 1592 CCTAAGGAGACTTGGTATGATTTAGACAGACTTCTGCTGGAAGAAATACAGATTTTCGG 1651
Db 467 TyrAsnGlnGluThrGluTyrMetGluGluValIleLeuGluLeuGlnIlePheLys 486
Qy 1652 GAACCGACTGCTATTTCAAGAAATGAGCTTTCCACTAAGCAGCAACAACTATATATGTT 1711
Db 487 AspProAlaProIleIleSerMetGluIleSerSerLysArgGlnGlnLeuTyrIleGly 506
Qy 1712 TCACCGCTGGGTGTCCTCCCTTACCGGTGTGATATTTACGGGAAGCGTGT 1771
Db 507 SerAlaSerAlaValAlaGlnValArgPheHisHisCysAspMetTyrGlySerAlaCys 526

Qy 1772 GCTGAGTGTTCCTCGCCCGAGACCTTACTGTCTGCTTGGGATGGTTCTGCATGTTCTCGC 1831
Db 527 AlaAspCysCysLeuAlaAlaArgAspProTyrCysAlaTrpAspGlyIleSerCysSerArg 546
Qy 1832 TATTTTCCACT-----GCAAAGAGAGCGCACAGACGACAGATATTAAGAAATGGA 1882
Db 547 TyrTyrProThrGlyAlaHisAlaLysArgArgPheArgArgGlnAspValArgHisGly 566
Qy 1883 GACCACCTGACTCTGTTACAGCTTACACCATGATAATCACCATTGGCCACAGCCCTCAA 1942
Db 567 AsnAlaAlaGlnGlnCysPheGlyGlnPheValGlyAspAlaLeuAspArgThrGlu 586
Qy 1943 GAGAAATCATCTATGTTGTAGAGATAGTAGACATTTTGGAAATGAGTCCGAAAGTCG 2002
Db 587 GluArgLeuAlaTyrGlyIleGluSerAsnSerThrLeuLeuGluCysThrProArgSer 606
Qy 2003 CAGAGAGCGTGTCTATTGGCAATTCAGAGGCGAAATGAAGAGCGGAAAGAGAGATC 2062
Db 607 LeuGlnAlaLysValIleTrpPheValGlnLysGlyArgAspValArgLysGluGluVal 626
Qy 2063 AGAGTGTGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAGTCTACAACAG 2122
Db 627 LysThrAspAspArgValValLysMetAspLeuGlyLeuLeuPheLeuArgValArgLys 646
Qy 2123 AAGGATTGAGCAATACCTCTGCTCGGTGGAACATGGGTTTCATACAACTCTTCTT 2182
Db 647 SerAspAlaGlyThrTyrPheCysGlnThrValGluHisAsnPheValHisThrValArg 666
Qy 2183 AAGTAACCTTGGAAAGTCATTCAGACAGAGCATTTGGAAGAACTTCTTCATAAAGATGAT 2242
Db 667 LysIleThrLeuGluValValGluGluHisLysValGluGlyMetPheHisLysAspHis 686
Qy 2243 GATGAGATGCTCTTAAGACCAAGAAATGTC-----AATAGCATGACACCT 2290
Db 687 GluGluGlu-----ArgHisHisLysMetProCysProLeuSerGlyMetSerGln 704
Qy 2291 AGCCAGAAAGTCTGTTAGACAGACTTCTCAGCTCATCACACCCCAATCTCAACACG 2350
Db 705 GlyThrLysProTyrTyrLysGluPheLeuGlnLeuIleGlyTyrSerAsnPheGlnArg 724
Qy 2351 ATGATGATGTTCTGTGAACAGTTTGGAAAAGGACCGAAGAACACGTCGCGCAAGGCCA 2410
Db 725 ValGluGluTyrCysGluLysValTyrCysThrAspLysLysArgLysLysLeuLysMet 744
Qy 2411 GGACATACCCAGGAAACAGTAACAATGGAAGCACTTACAGAAATTAAGAAAGGTAGA 2470
Db 745 -----SerProSerLysTrpLysTyrAlaAsnProGlnGluLys---Arg 758
Qy 2471 AACAGAGAGGACCCACGAAATTTGAGAGGCGCACCCAGG 2506
Db 759 LeuArgSerLysAlaGluHisPheArgLeuProArg 770

RESULT 10

US-08-556-422A-2
; Sequence 2, Application US/08556422A
; Patent No. 6576754
; GENERAL INFORMATION:
; APPLICANT: HALL, Kathryn T.
; APPLICANT: FREEMAN, Gordon J.
; APPLICANT: SCHULTZE, Joachim L.
; APPLICANT: BOUSSETIOTIS, Vassiliki
; APPLICANT: NADLER, Lee M.
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES
; FILE REFERENCE: DPN-005CPA2
; CURRENT APPLICATION NUMBER: US/08/556,422A
; CURRENT FILING DATE: 1995-11-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 862
; TYPE: PR1
; ORGANISM: Homo sapiens


```
Db 623 GlnCysLeuSerGluGluArgValIysAsnIysThrValPheGlnValAlaIalysHis 642
Qy 2189 -----ACCTCGAAGTCATTGAC 2205
Db 643 ValLeuGluValIysValProIysProValValAlaProThrLeuSerValValGln 662
Qy 2207 ACAGAG 2212
Db 663 ThrGlu 664

RESULT 11
US-08-556-422A-4
; Sequence 4, Application US/08556422A
; Patent No. 6576754
; GENERAL INFORMATION:
; APPLICANT: HALL, Kathryn T.
; APPLICANT: FREEMAN, Gordon J.
; APPLICANT: SCHULTZE, Joachim L.
; APPLICANT: BOUSSIOTIS, Vassiliki
; APPLICANT: NADLER, Lee M.
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES
; FILE REFERENCE: DFN-005CPA2
; CURRENT APPLICATION NUMBER: US/08/556,422A
; CURRENT FILING DATE: 1995-11-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 607
; TYPE: PRT
; ORGANISM: Mus musculus
US-08-556-422A-4

Alignment Scores:
Pred. No.: 9,7e-75 Length: 607
Score: 838.50 Matches: 210
Percent Similarity: 49.84% Conservative: 106
Best Local Similarity: 33.12% Mismatches: 249
Query Match: 17.13% Indels: 69
DB: 4 Gaps: 22

US-09-774-490-1 (1-2709) x US-08-556-422A-4 (1-607)
Qy 356 AACAGCTCCAGTATCATACCTCTCTTTGGATGAGAACGGAGTAGGCTGTATGTTGA 415
Db 12 AsnIleSerAsnTyrThrAlaLeuLeuSerGlnAspGlyLysThrLeuTyrValGly 31
Qy 416 GCAAAGGATCACATATTTTCATTCGACCTGTTAATATCAAG-----CAT 460
Db 32 AlaArgGluAlaLeuPheAlaLeuAsn---SerAsnLeuSerPheLeuProGlyGlyGlu 50
Qy 461 TTTCAAAGATGTGTGCCCATGATCTTTACACCCAGAGAGATGAATGCAAGTGGGCTGGA 520
Db 51 TyrGlnGluLeuLeuTyrSerAlaAspAlaAspArgLysGlnGlnCysSerPheLysGly 70
Qy 521 AAAGACATCTCGAAGAAATGTCTAATTTTCATCAAGTACTTAAGGCATATAATACAGACT 580
Db 71 LysAspProLysArgAspCysGlnAsnTyrIleLysIleLeuLeuProLeuAsnSerSer 90
Qy 581 CACTTGTACGCTGTGGACGGGGCTTTTCATCCATTTTCACCTACCTACATTAATGAATGGA 640
Db 91 HisLeuLeuThrCysGlyThrAlaAlaPheSerProLeuCysAlaTyrIleHisIleAla 110
Qy 641 CATCATCTCGAGGACAATATTTTAAAGCTGGAG-----AACTCACATTTT 685
Db 111 Ser-----PheThrLeuAlaGlnAspGluAlaGlyAsnValIleLeu 124
Qy 686 GAAACGCCCTGGGAAGATGCTCATATGACCTTAAGCTGTGACAGCATCCCTTTTAATA 745
Db 125 GluAspGlyLysGlyHisCysProPheAspProAsnPheLysSerThrAlaLeuValVal 144
Qy 746 GATGAGAAATATCTCTGGAACCTGCAGCTGATTTTATGGGGCAGACTTGTCTATCTTC 805
Db 145 AspGlyGluLeuTyrThrGlyThrValSerSerPheGlnGlnIysAsnAspProAlaIleSer 164
```

```
Qy 806 CGAACTCTTGGGCACCACCCCAATCAGGACAGACAGCATGATTCAGGTGGCTCAAT 865
Db 165 ArgSerGlnSer---SerArgProThrLysThrGluSer---SerLeuAsnTrpLeuGln 182
Qy 866 GATCAAAGTTCATTAGTGCCCACTCATCTCAGAGAGT---GACAATCCT---GAA 916
Db 183 AspProAlaPheValAlaSerAlaThrSerProGluSerLeuGlySerProIleGlyAsp 202
Qy 917 GATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGAGAAACACTCTGGA 976
Db 203 AsnAspLysIleTyrPhePheSerGluThrGlnGluPheGluPheGluAsn 222
Qy 977 GCTACTCACGCTAGATAGTGCATGATGCAAGAAATGACTTTTGGAGGSCACAGAACTCTG 1036
Db 223 ThrIleValSerArgValAlaArgValCysLysGlyAspGluGlyGlyGluArgValLeu 242
Qy 1037 GTGAATAAATGGACAACATCTCTCAAAAGCTCGTCTGATTTCTCAGTCCAGGTCCTCAAT 1096
Db 243 GlnGlnArgTrpThrSerPheLeuLysAlaGlnLeuLeuCysSerArgPro---AspAsp 261
Qy 1097 GGCATTGACACTCATTTTGTATGAATGCACTGAGATGATTTCTTAATG---AACTTTAAA 1150
Db 262 GlyPhePro-----PheAsnValLeuGlnAspValPheThrLeuAsnProAsnProGln 279
Qy 1151 GATCCTAAATCCAGTTGTATATGGAGTGTATTACGACTTCC-----AGTAACATTTTC 1204
Db 280 AspTrpArgLysThrLeuSerIleGlyValPheThrSerGlnTrpHisArgGlyThrThr 299
Qy 1205 AAGGGATCAGCGGTGTGTATGTATAGCATGATGATGTGAGAAGGGTTCCTTGGTGTCCA 1264
Db 300 GluGlySerAlaIleCysValPheThrMetAsnAspValGlnLysAlaPheAspGlyLeu 319
Qy 1265 TATGCCACACAGGATGACCAACTATCAATGGTGGCTTATCAAGGAAGAGTCCCTAT 1324
Db 320 TyrLysLysValAsnArgGluThrGlnGlnTrpTyrThrGluThrHisGlnValProThr 339
Qy 1325 CCACGGCCAGGAACCTTGT-----CCAGCAAAAACATTTGTGTGTTTTGACTCTACAAAG 1378
Db 340 ProArgProGlyAlaCysIleThrAsnSerAlaArgGluArgLysIleAsnSerSerLeu 359
Qy 1379 GACTTCTCATGATGATTTAATACCTTTTGAAGAAGTATCATCCAGCCATGTATACATCCAGTG 1438
Db 360 GlnLeuProAspArgValLeuAsnPheLeuLysAspHisPheLeuMetAspGlyGlnVal 379
Qy 1439 TTTCTATGAACATCCCAATAGTATCAAAACGGATGTAAATTTCAATTTACACAA 1498
Db 380 -----ArgSerArgLeuLeuLeuGlnProArgAlaArgTyrGln-----Arg 394
Qy 1499 ATTGCTCTAGACCCAGTGGATGACAGAGATGACAGATATGATGTTATGTTTATCGGAACA 1558
Db 395 ValAlaValHisArgValProGlyLeuHisSerThrTyrAspValLeuPheLeuGlyThr 414
Qy 1559 GATGTGGGACCGTCTTTAAAGTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAA 1618
Db 415 GlyAspGlyArgLeuHisLysAlaValThrLeuSerSerArgValHis-----430
Qy 1619 GAGTTCTGTGGAAGAATATCAGATTTTTCGGAACCGACTGCTATTTACAGCAATGGAG 1678
Db 431 -----IleIleGluGluLeuGlnIlePheProGlnGlnProValGlnAsnLeuLeu 448
Qy 1679 CTTTCCACTAAGCAGCAACACTATATATTGTTTCAACGCTGGGGTGTGCCAGCTCCCT 1738
Db 449 LeuAspSerHisGlyGlyLeuLeuTyrAlaSerSerHisSerGlyValValGlnValPro 468
Qy 1739 TTACACCGGTGTGATATTTACGGGAAGACCGTGTGCTGAGTGTGCTCGCCCGAGACCT 1798
Db 469 ValAlaAsnCysSerLeuTyr---ProThrCysGlyAspCysLeuLeuAlaArgAspPro 487
Qy 1799 TACTGTCTTGGATGGTCTGTCATGT-----TCTCGCTATTTTCCACTGCAAG 1849
Db 488 TyrCysAlaTrpThrGlySerAlaCysArgLeuAlaSerLeuTyrGlnProAspLeuAla 507
```


Qy	1073	ATTGTGCTCAGTCCAGGTCCAAATGGGCATTTGACACTCATTTTGAATGAACATGCAGGATGTA	1133
Db	314	AsnCysSerIleSerGlyGlu-----PheProPheTyrPheAsnGluIleGlnSerVal	331
Qy	1133	TTCCCTAATGAACATTAAAGATCTCAAATAATCCAGTGTGTATATGGAGTGTTTACGACTTCC	1192
Db	332	TyrGlnLeuProSerAspLysSerArg-----PhePheAlaThrPheThrThrSer	348
Qy	1193	AGTAACACATTTTCAAGGGATCAGCCGTGTGTATATAGCATGAGTGTATGGAGAGGGGTG	1252
Db	349	ThrAsnGlyLeuIleGlySerAlaValCysSerPheHisIleAsnGluIleGlnAlaAa	368
Qy	1253	TTCCCTTGGTCATATGCCACAGGAGTGGACCAACTATCAATGGGTGCTC---TATCAA	1309
Db	369	PheAsnGlyLysPheLysGluGlnSerSerSerAsnSerAlaTrpLeuProValLeuAsn	388
Qy	1310	GGAAAGTGCCTTATCCACGCGCAGGAACTGTGCCAGCAAAACATTTGTGTGTTTGAC	1369
Db	389	SerArgValProGluProArgProGlyThrCys-----ValAsn	401
Qy	1370	TCTACAAAGGACCTTCTCGATGATGTTTAACTTTGCAAGAAAGTCATCCAGGCATGTAC	1429
Db	402	AspThrSerAsnLeuProAspThrValLeuAsnPheIleArgSerHisProLeuMetAsp	421
Qy	1430	AATCCAGTGTTCCTATGAACAATCGCCCAATAGTGTATCAAAACGGATGTAATATCAA	1489
Db	422	LysAlaValAsnHisGluHisAsnAsnProValTyrTyrLysArgAspLeu-----Val	439
Qy	1490	TTTACACAATTTGTCGTAGAC-----CGAGTGGATCCAGAAGTGCACAGTATGATGTT	1543
Db	440	PheThrLysLeuValValAspLysIleArgIleAspIleLeuAsnGlnGluTyrIleVal	459
Qy	1544	ATGTTTATCGAAACAGATGTTGGACCGTCTCTTAAAGTAGTGTTCATTCCTAAGGAGACT	1603
Db	460	TyrTyrValGlyThrAsnLeuGlyArgIleTyrLysIleValGln-----	474
Qy	1604	TGGTATGATTTAGAAGAGTTCTGCTGGAGAAGATGACAGTGTTCGG-----GAACCG	1657
Db	475	TyrTyrArgAsnGlyGluSerLeuSerLysLeuLeuAspIlePheGluValAlaProAsn	494
Qy	1658	ACTGCTATTTCCAGCAATGGAGCTTTCACATAGCAGCAACACTATATATATTTGTTCAACG	1717
Db	495	GluAlaIleGlnValMetGluIleSerGlnThrArgLysSerLeuTyrIleGlyThrAsp	514
Qy	1718	GCTGGGTGTCCAGCTCCCTTTACACCGGTGTGATATTTACGGGAAGCGTGTCTGTGAG	1777
Db	515	HisArgIleLysGlnIleAspLeuAlaMetCysAsnArgArgTyrAspAsnCysPheArg	534
Qy	1778	TGTTGCTCCGCCGAGACCTTACTGTGTGGATGGTGTCTGCA-----TGTTCTCGC	1831
Db	535	Cys-----ValArgAspProTyrCysGlyTyrAspLysGluAlaAsnThrCysArgPro	552
Qy	1832	TATTTTCCCACTGCAAGAGACGCACAGACGACAGATATAGAAATGGAGACCCACTG	1891
Db	553	Tyr-----GluLeuAspLeuLeuGlnAspValAlaAsn-----	563
Qy	1892	ACTCACTGTTTCAGACTTACACCATGATAATCACCATGGCCACAGCCCTGAAGAGAGAATC	1951
Db	564	--GluThrSerAspIleCysAspSerSer-----ValLeuLysLysLysIle	578
Qy	1952	ATC-----TATGGTGTAGAGATAGTAGCACAATTTTGGAAATGCAGTCCGAGTCCGAC	2005
Db	579	ValValThrTyrGly-----GlnSerValHisLeuGlyCysPheValLysIlePro	595
Qy	2006	AGAGCGCTG-----GTCTATTGGCAATTCAGAGGGCGAAATGAAGACGAAAA	2053
Db	596	GluValLeuLysAsnGluInValThrTrp---TyrHisHisSerLysAspLysGlyArg	614
Qy	2054	GAAGAGATCAGATGGAT-----GATCATATATCATCAGGACAGATCAAGGCCCTTCGCTA	2107
Db	615	TyrGluIleArgTyrSerProThrLysTyrIleGluThrThrGluArgGlyLeuValVal	634
Qy	2108	CGTAGTCTACACAGAGAGGATTGAGCAATTAAGTCTGTCAT	2149

[illegible]

Db 54 GlyAlaAspHisValArgGluPheAsn-----CysGlyLysLeuTyrTyrArgThrPhe 71
QY 380 CTTTGGATGAGGAACGAGTATGCTGTATGTTGGAGCAAGGATCAATATTTTCATTC 439
Db 72 HisMetAsnGluAspArgAspThrLeuTyrValGlyAlaMetAspArgValPheArgVal 91
QY 440 GACCTGGTTATATC-----AAGGATTTCAAAGATTTGTGTGCCAGTATCTTACACC 493
Db 92 AsnLeuGlnAsnIleSerSerSerAsnCysAsnArgAspAlaIleAsnLeuGluProThr 111
QY 494 AGAAGAGAT-----GAATGCGAAGTGGCTGGAAGAGACATCTCAAAAGATGCTCAAT 547
Db 112 ArgAspAspValValSerCysValSerLysGlyLysSerGlnIlePheAspCysLysAsn 131
QY 548 TTCATCAAGGTACTTAAGCATATAATCAG---ACTCACTTGATGACGCTGTGGAAACGGGG 604
Db 132 HisValArgValIleGlnSerMetAspGlnGlyAspArgLeuTyrValCysGlyThrAsn 151
QY 605 GCTTTTCATCAATTTGCACCTACATTTGAAATGGACATCATCTCGAGGAC----- 655
Db 152 Ala-----HisAsnProLysAspTyrValIle 160
QY 656 -----AATATTTTAAGCTGGAGAACTCACATTTTGAAGCGCGCTGG----- 700
Db 161 TyrAlaAsnLeuThrHisLeuProArgSerGluTyrValIleGlyValGlyLeuGlyIle 180
QY 701 ---AAGAGTCCATATGACCTTAAGCTGTGACAGCATCCCTTTTAAATAGATGA----- 751
Db 181 AlaLysCysProTyrAspProLeuAspAsnSerThrAlaIleTyrValGluAsnGlyAsn 200
QY 752 -----GAATTATCTCGAAGTGCAGCTGATTTTATGGGCGAGACTTT 796
Db 201 ProGlyGlyLeuProGlyLeuTyrSerGlyThrAsnAlaGluPheThrLysAlaAspThr 220
QY 797 GCTATCTTCGGAACCTCTGGGACCAACACCAATC----- 832
Db 221 ValIlePheArgThrAspLeuTyrAsnThrSerAlaLysArgLeuGluTyrLysPheLys 240
QY 833 AGGACAGACGACGATGATTCAGGTGGCTCAATGATCCAAAGTTCATTTAGTGGCCACCTC 892
Db 241 ArgThrLeuLysTyrAspSerLysTrpLeuAspLysProAsnPheValGlySerPheAsp 260
QY 893 ATCTCAGAGAGTGACAATCTCTGAAGTACAAAGTATCTTTCTCCGTGAAAATGCA 952
Db 261 IleGlyGlu-----TyrValTyrPhePheArgGluThrAla 273
QY 953 ATAGATGGAGAACACTCTGGAAGCTACTACGCTAGATAGGTGCAGATATGCAAGAT 1012
Db 274 ValGluTyrIleAsnCysGlyLysAlaValTyrSerArgIleAlaArgValCysLysLys 293
QY 1013 GACTTTGAGGGCAGAGAGTCTGTGTGAATAATGGACAACATTCCTCAAGCTCGTCTG 1072
Db 294 AspValGlyLysAsnLeuLeuAlaHisAsnTrpAlaThrTyrLeuLysAlaArgLeu 313
QY 1073 ATTTCTCAGTCCAGGTCCTCAATGGCATTGACATCTATTTGTATGTAAGTCAAGTATGTA 1132
Db 314 AsnCysSerIleSerGlyGlu-----PheProPheTyrPheAsnGluIleGlnSerVal 331
QY 1133 TTCCTAATGAACCTTTAAAGATCCTAAAAATCCAGTTGTATGGAGTGTTCACGACTTCC 1192
Db 332 TyrGlnLeuProSerAspLysSerArg-----PhePheAlaThrPheThrThrSer 348
QY 1193 AGTAACATTTTCAAGGATCAGCGGTGTGTATGATAGCATGAGTGTGAGAGGGTG 1252
Db 349 ThrAsnGlyLeuIleGlySerAlaValCysSerPheHisIleAsnGluIleGlnAlaAla 368
QY 1253 TTCCTTGGTCCATATGCCCCACAGGATGGACCAACTATCAATGATGGTGCCT---TATCAA 1309
Db 369 PheAsnGlyLysPheLysGluGlnSerSerAsnSerAlaTrpLeuProValLeuAsn 388
QY 1310 GGAAGATGCCCTATCCAGCGCCAGGAACCTTGTCCCGCAAAACATTTGGTGTGTTTGC 1369
;

Db 389 SerArgValProGluProArgProGlyThrCys-----ValAsn 401
QY 1370 TCTACAAGGACCTTCCTGATGATGTTATATACTTTCACAAGTTCATCCAGCCATGTAC 1429
Db 402 AspThrSerAsnLeuProAspThrValLeuAsnPheIleArgSerHisProLeuMetAsp 421
QY 1430 AATCCAGTGTTCATAGCAACATCGCCCAATAGTAGTCAAAAGCGGATGTAAATTTATCAA 1489
Db 422 LysAlaValAsnHisGluHisAsnAsnProValTyrTyrLysArgAspLeu-----Val 439
QY 1490 TTTACACAAATTTGCTAGAC-----CGAGTGGATGCAGAGATGGACAGTATGTT 1543
Db 440 PheThrLysLeuValValAspLysIleArgIleAspIleLeuAsnGlnGluTyrIleVal 459
QY 1544 ATGTTTATCGAAACAGATGTTGGACCGTCTTAAAGTAGTTCNAATTCCTAAGGAGACT 1603
Db 460 TyrTyrValGlyThrAsnLeuGlyArgIleTyrLysIleValGln----- 474
QY 1604 TGGTATGATTTAGAAGAGAGTTCCTGCGAAGAAATGCAGATTTTTCGG-----GAACCG 1657
Db 475 TyrTyrArgAsnGlyGluSerLeuSerLysLeuAspIlePheGluValAlaProAsn 494
QY 1658 ACTCTATTTCAGCAATGGAGCTTTCCTACTAAGCAGCAACACTATATATTTGGTTCAACG 1717
Db 495 GluAlaIleGlnValMetGluIleSerGlnThrArgLysSerLeuTyrIleGlyThrAsp 514
QY 1718 GCTGGGTGTCACAGCTCCCTTTACCGGTGTGATATTTACGGGAAAGCTGTGCTGAG 1777
Db 515 HisArgIleLysGlnIleAspLeuAlaMetCysAsnArgArgTyrAspAsnCysPheArg 534
QY 1778 TGTTCCTCGCCCGAGACCTTACTGTCTTGGGATGGTCTGCA-----TGTTCCTCGC 1831
Db 535 Cys-----ValArgAspProTyrCysGlyTrpAspLysGluAlaAsnThrCysArgPro 552
QY 1832 TATTTTCCACTGCTCAAGAGACGACAGACGACAGATATAGAAATGAGAGCCCACTG 1891
Db 553 Tyr-----GluLeuAspLeuGlnAspValAlaAsn----- 563
QY 1892 ACTCACTGTTTCAGACTTACACCATGATAATCACCATGCGCCACGCTGNAAGAGAGATC 1951
Db 564 ---GluThrSerAspIleCysAspSer-----ValLeuLysLysLysIle 578
QY 1952 ATC-----TATGCTGTAGAGATAGTAGACATTTTGGTAAATGACGTCCGAAAGTCGAG 2005
Db 579 ValValThrTyrGly-----GlnSerValHisLeuGlyCysPheValLysIlePro 595
QY 2006 AGACGCTG-----GTCTATTGGCAATTCAGAGCGCAATGAGAGCGCAAAA 2053
Db 596 GluValLeuLysAsnGluGlnValThrTrp---TyrHisHisSerLysAspLysGlyArg 614
QY 2054 GAAGAGATCAGATGGAT-----GATCATATCATCAGACAGATCAAGGCTTCTGCTA 2107
Db 615 TyrGluIleArgTyrSerProThrTyrIleGluThrThrGluArgGlyLeuValVal 634
QY 2108 CGTAGTCTCAACAAAGAGGATTTCAGCAATTTACCTCTGCGCAT 2149
Db 635 ValSerValAsnGluAlaAspGlyArgTyrAspCysHis 648

RESULT 14

US-09-060-692-62
; Sequence 62, Application US/09060692
; Patent No. 5935865
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200

CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/121,713
FILING DATE: 13-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:


```
Db 221 ValIlePheArgThrAspLeuTyrAsnThrSerAlaLysArgLeuGluTyrLysPheLys 240
QY 833 AGGACAGACGATGATTCAGGTGGCTCAATGATCCAAAGTTCATTAGTCCACACCTC 892
Db 241 ArgThrLeuLysTyrAspSerLysIlePheAspLysProAsnPheValGlySerPheAsp 260
QY 893 ATCTCAGAGAGTACAAATCCTGAAGATGACAAAGTATATCTTTCTTCGCGTGAATGCA 952
Db 261 IleGlyGlu-----TyrValTyrPhePhePheArgGluThrAla 273
QY 953 ATAGATGGAGACACTCTGGAAGCTACTCAGCTAGATAGTTCAGATATGCAAGAT 1012
Db 274 ValGluTyrIleAsnGlyLysAlaValTyrSerArgIleAlaArgValCysLysLys 293
QY 1013 GACTTTGGAGGCGACAGAAGTCTGGTGAATTAATGGACAACATCTCTCAAGCTCGTCTG 1072
Db 294 AspValGlyGlyLysAsnLeuLeuAlaHisAsnTTPAlaThrTyrLeuLysAlaArgLeu 313
QY 1073 ATTTCCTCAGTCCAGGTCCTCAATGGCATTGACACTCATTTTGTAGTAAGTGCAGATGTA 1132
Db 314 AsnCysSerIleSerGlyGlu-----PheProPheTyrPheAsnGluIleGlnSerVal 331
QY 1133 TTCCTAATGAACCTTTAAGATCCTAAATCCAGTTGTATATGGAGTGTTCAGACTTCC 1192
Db 332 TyrGlnLeuProSerAspLysSerArg-----PhePheAlaThrPheThrThrSer 348
QY 1193 AGTAACATTTTCAAGCGATCAGCGCTGTGTATGATGATGAGTGTGTGAGAAGGGTG 1252
Db 349 ThrAsnGlyLeuIleGlySerAlaValCysSerPheHisIleAsnGluIleGlnAlaAla 368
QY 1253 TCTCTTGGTCCATATGCCACAGGATGAGACCAACTATCAATGGTGCCT---TATCAA 1309
Db 369 PheAsnGlyLysPheLysGluGlnSerSerAsnSerAlaTyrLeuProValLeuAsn 388
QY 1310 GGAAGAGTCCCTATCCACGGCCAGGAAGTGTCTCCAGCAAAACATTTGGTGGTTTCAC 1369
Db 389 SerArgValProGluProArgProGlyThrCys-----ValAsn 401
QY 1370 TCTACAAAGGACCTTCTCGATGATGTATTAACCTTTGCAAGAAGTATCCAGCCATGTAC 1429
Db 402 AspThrSerAsnLeuProAspThrValLeuAsnPheIleArgSerHisProLeuMetAsp 421
QY 1430 AATCCAGTGTTCCTATGAACAATCGCCCAATAGTATGATCAAAACCGATGTAAATATCAA 1489
Db 422 LysAlaValAsnHisGluHisAsnAsnProValTyrTyrLysArgAspLeu-----Val 439
QY 1490 TTTACACAAATGTCGTAGAC-----CGAGTGGATGCAGAGATGGACAGTATGATGTT 1543
Db 440 PheThrLysLeuValValAspLysIleArgIleAspIleLeuAsnGlnGluTyrIleVal 459
QY 1544 ATGTTTATCGGAACAGATGTTGGGACCGTCTTAAAGTAGTTTCAATTCTTAAGGAGACT 1603
Db 460 TyrTyrValGlyThrAsnLeuGlyArgIleTyrLysIleValGln-----474
QY 1604 TGGTATGATTTAGAGAGGTTCTCTGGAAGAAATACACAGTTCCTTCGG-----GAACCG 1657
Db 475 TyrTyrArgAsnGlyGluSerLeuSerLysLeuLeuAspIlePheGluValAlaProAsn 494
QY 1658 ACTGCTATTTACGCAATGAGCTTCTCCTAAGCAGCAACACTATATATTTGGTTCAACG 1717
Db 495 GluAlaIleGlnValMetGluIleSerGlnThrArgLysSerLeuTyrIleGlyThrAsp 514
QY 1718 GCTGGGGTTCGCCAGCTCCCTTTACACCGGTGTGATATTACGGGAAGCGTGTGCTGAG 1777
Db 515 HisArgIleLysGlnIleAspLeuAlaMetCysAsnArgArgTyrAspAsnCysPheArg 534
QY 1778 TGTTCCTCGCCCGGACAGCCTTACTGTGCTTGGGATGGTCTCTGCA-----TGTTCCTCGC 1831
Db 535 Cys-----ValArgAspProTyrCysGlyTyrAspLysGluAlaAsnThrCysArgPro 552
QY 1832 TATTTTCCCACTGCAAGAGCGCACAGACGACAGATATATAAGAAATGGAGCCCACTG 1891
```

```
Db 553 Tyr-----GluLeuAspLeuLeuGlnAspValAlaAsn-----563
QY 1892 ACTCACTGTTTACAGACTTACACCATGATTAATCACCATGGCCACAGCCCTGAAGAGAGAATC 1951
Db 564 ---GluThrSerAspIleCysAspSerSer-----ValLeuLysLysLysIle 578
QY 1952 ATC-----TATGGGTAGAGAATAGTAGACATTTTGGAAATGCAAGTCCGAGTCCGAG 2005
Db 579 ValValThrTyrGly-----GlnSerValHisLeuGlyCysPheValLysIlePro 595
QY 2006 AGAGCGCTG-----GTCTATTGGCAATTCCAGAGCGCAAAATCAAGAGCGAAAA 2053
Db 596 GluValLeuLysAsnGluGlnValThrTrp---TyrHisHisSerLysAspLysGlyArg 614
QY 2054 GAAGAGATCAGAGTGGAT-----GATCATATCATCAGACAGATCAAGGCCTTCTGTCTA 2107
Db 615 TyrGluIleArgTyrSerProThrLysTyrIleGluThrThrGluArgGlyLeuValVal 634
QY 2108 CGTAGTCTTACAAACAGAGAGATTCAGGCAATTACCTCTGCCAT 2149
Db 635 ValSerValAsnGluAlaAspGlyGlyArgTyrAspCysHis 648
```

Search completed: August 3, 2003, 10:20:43
Job time : 64 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 10:29:36 ; Search time 50.5 Seconds
(without alignments)
10317.662 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 895

Sequence: 1 aatctttattttatcatgatg.....agggtttttttcctaataacc 2709

Scoring table:

OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283308 seqs, 96168682 residues

Word size: 1

Total number of hits satisfying chosen parameters: 565704

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09774490/runat_03082003.102927.4471/app.query.fasta_1.2887
-DB=PIR_76 -QFMT=fastan -SUFFIX=oligo.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=BITS -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi -LIST=100
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=75 -MOD=LOCAL -OUTFMT=ptc
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09774490 @CGN 1.1.69 @runat_03082003.102927.4471 -NCPU=6 -ICPU=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	771	86.1	771	2 D49423	semaphorin III pre
2	82	9.2	666	2 I58169	semaphorin III - m
3	79	8.8	772	2 I48747	semaphorin D - mou
4	58	6.5	772	2 A49069	collapsin - chickc
5	20	2.2	748	2 I48744	semaphorin A - mou
6	15	1.7	749	2 G01856	semaphorin V - hum
7	15	1.7	751	2 I48748	semaphorin E - mou
8	13	1.5	753	2 G02173	semaphorin III fam
9	10	1.1	653	2 T03102	semaphorin homolog
10	10	1.1	834	2 S66498	M-sema F protein p
11	9	1.0	311	2 T11017	MLCL536.34 protein
12	9	1.0	688	1 C1HUS	complement subcomp
13	9	1.0	782	2 I48746	semaphorin C - mou
14	9	1.0	1244	2 S25327	cytoskeleton assem

ryanodine receptor	9	1.0	4859	2	S74173
ryanodine-binding	9	1.0	4868	2	B54161
ryanodine receptor	9	1.0	4869	2	S66572
ryanodine receptor	9	1.0	4872	2	S27272
ryanodine receptor	9	1.0	4967	2	S72269
ryanodine receptor	9	1.0	4969	2	A37113
ryanodine-binding	9	1.0	5037	1	A54161
ryanodine receptor	9	1.0	5126	2	S40450
pnd protein - Esch	8	0.9	50	1	S02763
tryptophan biosynt	8	0.9	50	2	JU0071
hypothetical prote	8	0.9	75	2	B38256
hypothetical prote	8	0.9	101	2	C71047
butyrate-acetoacet	8	0.9	178	2	AG0244
probable endociti	8	0.9	223	2	B83932
hypothetical prote	8	0.9	283	2	D84868
oligopeptide ABC t	8	0.9	309	2	G87498
oligopeptide ABC t	8	0.9	326	2	H97346
spore germination	8	0.9	350	2	C97346
proteinase inhibit	8	0.9	368	1	I39856
G protein-coupled	8	0.9	374	2	A59273
type I site-specif	8	0.9	377	2	B30341
hypothetical prote	8	0.9	409	2	T09461
hypothetical 53.0K	8	0.9	421	2	T22969
probable integral	8	0.9	463	2	J50376
serine-type D-Ala-	8	0.9	475	2	G95379
B. subtilis negati	8	0.9	479	2	A64117
B. subtilis negati	8	0.9	571	2	AB1274
D-lactate dehydrog	8	0.9	571	2	AC1637
hypothetical prote	8	0.9	579	2	S51528
phosphate acetyltr	8	0.9	697	2	T37827
hypothetical prote	8	0.9	713	2	S50130
semaphorin B - mou	8	0.9	745	2	T33751
hypothetical prote	8	0.9	760	2	I48745
probable protein A	8	0.9	817	2	T51787
semaphorin alpha c	8	0.9	941	2	A86404
semaphorin F precu	8	0.9	966	2	A33626
probable ABC trans	8	0.9	1074	2	JC5928
hypothetical prote	8	0.9	1146	2	F84487
ryanodine receptor	8	0.9	2420	2	A84652
ryanodine receptor	8	0.9	5032	1	A35041
ryanodine receptor	8	0.9	5035	1	I46646
phycobiliprotein s	8	0.9	5037	2	B35041
interleukin-7 rece	8	0.9	10797	2	T30192
hypothetical prote	8	0.8	16	2	S09700
photosystem I prot	7	0.8	22	2	B40256
alpha-amylase (EC	7	0.8	35	2	B82805
hypothetical prote	7	0.8	37	2	S03674
hypothetical prote	7	0.8	51	2	S05490
hypothetical prote	7	0.8	63	2	D42194
hypothetical prote	7	0.8	68	2	S33581
hypothetical prote	7	0.8	87	2	A83261
hypothetical prote	7	0.8	88	2	T25449
hypothetical prote	7	0.8	88	2	T30626
hemoglobin alpha c	7	0.8	89	2	I68530
hypothetical prote	7	0.8	91	2	AF2068
potassium-dependen	7	0.8	98	2	AE2336
hypothetical prote	7	0.8	99	2	G84410
hypothetical prote	7	0.8	99	2	A69514
hypothetical prote	7	0.8	100	2	AD2566
conserved hypotnet	7	0.8	102	2	D70408
insulin if precurs	7	0.8	106	1	IPXL2
hypothetical prote	7	0.8	106	2	T47814
conserved hypotnet	7	0.8	106	2	H82421
hypothetical prote	7	0.8	106	2	B64001
hypothetical prote	7	0.8	107	2	G72496
hypothetical prote	7	0.8	110	2	C70599
hypothetical prote	7	0.8	111	2	T00688
hypothetical 14K p	7	0.8	118	2	E33958
probable integral	7	0.8	119	2	F81894
crsB protein NMB10	7	0.8	119	2	C81125
conserved hypotnet	7	0.8	122	2	F75316
hypothetical prote	7	0.8	131	2	E75350
hypothetical prote	7	0.8	131	2	E72469

88 7 0.8 132 2 D72151 B12L protein - var
c 89 7 0.8 135 2 E81433 probable integral
c 90 7 0.8 138 2 D72727 hypothetical prote
c 91 7 0.8 138 2 T18449 pathogenicity fact
c 92 7 0.8 141 2 A32268 carcinomembryonic a
c 93 7 0.8 141 2 I57007 pregnancy-specific
c 94 7 0.8 141 2 C24338 hemoglobin alpha-T
c 95 7 0.8 141 2 T10195 hypothetical prote
c 96 7 0.8 142 2 S71107 cell-cell adhesion
c 97 7 0.8 145 2 T00987 hypothetical prote
c 98 7 0.8 145 2 E90398 hypothetical prote
c 99 7 0.8 147 2 D72504 hypothetical prote
c 100 7 0.8 147 2 A84546 50S ribosomal prot

ALIGNMENTS

RESULT 1
D49423
semaphorin III precursor - human
C:Species: Homo sapiens (man)
C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 24-Sep-1999
C:Accession: D49423
R:Kolodkin, A.L.; Matthes, D.J.; Goodman, C.S.
Cell 75, 1369-1399, 1993
A:Title: The Semaphorin genes encode a family of transmembrane and secreted growth cone
A:Reference number: A49423; MUID:94094332; PMID:8269517
A:Accession: D49423
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-771 <KOL>
A:Cross-references: GB:L26081; NID:9799328; PIDN:AAA65938.1; PID:g436560
C:Genetics:
A:Gene: GDB:SEMA1
A:Cross-references: GDB:283448
C:Superfamily: semaphorin

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x D49423 (1-771)

QY 200 ATGGGCTGGTTAACTAGGATTCCTCTCTTTCTGGGGAGTATTACTTACAGCAAGAGCA 259
DB 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGATGGGAAGAACAAATGTGCCAGGCTGAAATATCTCTCAAGAAAGATGTTG 319
DB 21 AsnTyrGlnAsnGlyLysAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAATGTGATCACTTTCAATGCTGGCTGGCCCAAGAGTCCAGTATCATACCTTC 379
DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
QY 380 CTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAGAGATCAATATTTTCATTC 439
DB 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisLeuPheSerPhe 80
QY 440 GACCTGGTTATATCAAGATTTTCAAAAGATGTGTGGCCAGTATCTTACACCAAGAGA 499
DB 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATGAATCAAGTGGCTGGAAAGACATCTCTGAAGATGTGCTAAATTTTCATCAAGTA 559
DB 101 AspGluCysLysIleTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheLysVal 120
QY 560 CTTAAGGCATATATCAGACTCATTGTACGCTGTGGAGCGGGGCTTTTCATCCCAATT 619

Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyValaPheHisProIle 140
QY 620 TGCACCTACATTGAATTTGGACATCATCTCAGAGACATATTTTAACTCGAGACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAAAACGGCCGTGGGAAGAGTCCATATGACCCCTAAGCTCTCAGCAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGAGAAATATATCTCTGGAACCTCAGCTGATTTTATGGGCGAGACTTGTCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCCGAATCTTGGCCACACACCCCATCAGGACAGACAGCAGCATGATCCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCAATGATCAAAAGTTTCATTAGTCCACCTCATCTCAGAGAGTGACAAATCCTCAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTTCTCCGTGAAATCAATAGATGAGAACACATCTCGAAAGACT 979
Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGATAGTGTAGATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGGCAACATCTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGAGTCCCAATGGC 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTATTTGATGAACACTCAGAGATGATTTCTTAATGAATTTAAAGATCCTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGTATGAGTGTGTTTACGACTTCCAGTAACTTTTCAAGGATCAGCCGTG 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGATGTGAGAAAGGTGTCTTCTTGTGCTCATATGCCACAGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCAACTATCAATCGGTGCTTATCAAGGAAGAGTCCCTATCCACGCCAGGAAT 1339
Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCAGCAAAACATTTGGTGGTTTGTGCTCTCAAGAGACCTTCTCTGATGATGTTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAAGAAGTCACTCCAGCCATGTACAATCCAGTGTCTTCTATGAAACAATCGCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACGGATGTAATTTATCAATTTTACAAATTTGCTAGACCGAGTGGAT 1519
Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAATGCAAGTATGATGTTATGTTTATCGGAACAGATGTGGGACCGTCTCTTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAGAGGTTCCTGCTCGAAGAAATG 1639
Db 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTTCGGGAACCGACTGTCTATTTTCAGCAATGAGGCTTTTCACCTAACGACAGCAA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500

QY 1700 CTATATATTGTTCAACGGCTGGGTTCCCGAGTCTCCCTTTACACCGGTGATATTATAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGCGTGTGCTGAGTGTGCTCCCGAGACCCCTTACTGTGCTGGATGGTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTCCCACTCCAAAGAGACGACACAGACGACAGATATAGAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgThrArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTCATCTGTTAGATTAACCATGATATATCACTAGCCACAGCCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisAspAsnHisGlyHisSerPro 580
QY 1940 GAAGAGAGAACNTATGTTGTAGAGATAGTACACATTTTGGATGAGTCCGAG 1999
Db 581 GluGluArgIleIleTyrGlyValGluAsnSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGCGCTGCTTATTGGCAATTCAGAGCGCAATGAAGAGCGAAAGAGAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGAGTGGATGATCATATCATCAGGACAGATCAAGCCCTTCTGCTACGTAGTCTCAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGGATTCAGCAATTACTCTGCGCATCGGTGGAAACATGGTTTCATACAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAAACCTGGAAGTCAATGACACAGAGCAATTTGGAAGAACTTCTTCATAAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLysAsp 680
QY 2240 GATGATGAGATGCTCTTAAGACCAAGAAATGTCCTCAATAGCATGACACCTAGCCAGAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGGTACAGACTTCATGCTGCTCATCAACCCCAATCTCAACACATGATGATGAG 2359
Db 701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAAGTTTGAAGAGGACCGAAGCAACAGCTCG3CAAGGCCAGGACATACC 2419
Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgArgGlnArgProGlyHisThr 740
QY 2420 CCAGGAAACAGTAACAATGGAAGCATTACAAGAAATAAGAAAGGTAGAAACAGGAGG 2479
Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
QY 2480 ACCACGAAATTGAGAGGCGACCCAGGAGTGTC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771
RESULT 2
158169
semaphorin III - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 24-Sep-1999
C:Accession: I58169
R:Messersmith, E.K.; Leonardo, E.D.; Shatz, C.J.; Tessier-Lavigne, M.; Goodman, C.S.; Kie
Neuron 14, 949-959, 1995
A:Title: Semaphorin III can function as a selective chemorepellent to pattern sensory pr
A:Reference number: 158169; MUID:95267432; PMID:7748562
A:Accession: I58169
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-666 <RES>
A:Cross-references: GB:L40484; NID:g703189; PIDN:AAA73934.1; PID:g703190
C:Genetics:

A:Gene: Semaphorin III
C:Superfamily: semaphorin
Alignment Scores:
Pred. No.: 5.44e-79 Length: 666
Score: 82.00 Matches: 82
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.16% Indels: 0
DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x I58169 (1-666)
QY 999 GAGAGTCAATCTGAGATGACAAAGTATATCTTTTCCGTGAAAATGCAATAGAT 958
Db 128 GluSerAspAsnProGluAspAspLysValTyrPhePheArgGluAsnAlaIleAsp 147
QY 959 GGAGAACACTCTGAAAAGCTACTCACGCTAGATAGTTCAGATATGCAAGATGACTTT 1018
Db 148 GlyGluHisSerGlyLysAlaThrHisAlaArgIleGlyGlnIleCysLysAsnAspPhe 167
QY 1019 GGAGGACACAGACTCTGGTGAATAAATGGAACAACATTTCTCAAGCTCGTCTGATTGC 1078
Db 168 GlyGlyHisArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeuIleCys 187
QY 1079 TCACTGCCAGTCCAAATGCGATTGACACTCATTTTGTATGAATGCGAGATGATTCTTA 1138
Db 188 SerValProGlyProAsnGlyIleAspThrHisPheAspGluLeuGlnAspValPheLeu 207
QY 1139 ATGAAC 1144
Db 208 MetAsn 209
RESULT 3
148747
semaphorin D - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C:Accession: I48747
R:Puschel, A.W.; Adams, R.H.; Betz, H.
Neuron 14, 941-948, 1995
A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates
A:Reference number: 148744; MUID:95267431; PMID:7748561
A:Accession: I48747
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-772 <RES>
A:Cross-references: EMBL:X85993; NID:g854329; PIDN:CAA59985.1; PID:g854330
C:Genetics:
A:Gene: semD
C:Superfamily: semaphorin
Alignment Scores:
Pred. No.: 9.21e-76 Length: 772
Score: 79.00 Matches: 79
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.83% Indels: 0
DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x I48747 (1-772)
QY 1169 GTATATGAGTGTATTACACTTCCAGTAAACATTTTCAAGGATCAGCGTGTATGTAT 1228
Db 324 ValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaValCysMetTyr 343
QY 1229 AGCATGATGATGTGAGAAGGGTTCCTTGTGTCATATGCCACAGGATGAGACCCAAAC 1288
Db 344 SerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAspGlyProAsn 363
QY 1289 TATCAATGGTGCCTTATCAAGAGAGTCCCTATCCAGCCGAGCACTTGTCGCCAGC 1348
Db 364 TyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThrCysProSer 383

QY 1349 AAAACATTTGGTGGTTTCTGACCTACAAAGGACCTTCTCGATGATGCTTATACCTTT 1405
|||||
Db 384 LysThrPheGlyCysPheAspSerThrLysAspLeuProAspValIleThrPhe 402
|||||

RESULT 4
A49069
semaphorin - chicken
C:Species: Gallus gallus (chicken)
C>Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 24-Sep-1999
C:Accession: A49069
Cell 75, 217-227, 1993
A:Title: Collapsin: a protein in brain that induces the collapse and paralysis of neuron
A:Reference number: A49069; MUID:94006554; PMID:9402908
A:Accession: A49069
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-772 <LUO>
A:Cross-references: GB:U02528; NID:9410078; PIDN:AAC59638.1; PID:g410079
C:Superfamily: semaphorin

Alignment Scores:
Pred. No.: 4.09e-53 Length: 772
Score: 58.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.48% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x A49069 (1-772)

QY 971 GGAAGAGTACTACGCTAGATAGTCAGATATCAAGATGACTTTGGAGGCACAGA 1030
|||||
Db 258 GlyLysAlaThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArg 277
|||||

QY 1031 AGCTCGGTGAATAAATGACACATTCCTCAAGCTCGTCTGATTTGCTCAGTGCAGGT 1090
|||||
Db 278 SerLeuValAsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGly 297
|||||

QY 1091 CCAATGGCATGACACTCATTTTGTATGATGACCTGACAGGATGATTCCTTAATGAAC 1144
|||||
Db 298 ProAsnGlyIleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsn 315
|||||

RESULT 5
I48744
semaphorin A - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C:Accession: I48744
R:Busch, A.W.; Adams, R.H.; Betz, H.
Neuron 14, 941-948, 1995
A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates
A:Reference number: I48744; MUID:95267431; PMID:7748561
A:Accession: I48744
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-748 <RES>
A:Cross-references: EMBL:X85990; NID:9854323; PIDN:CAA59982.1; PID:9854324
C:Genetics:
A:Gene: sema
C:Superfamily: semaphorin

Alignment Scores:
Pred. No.: 3.94e-12 Length: 748
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.23% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x I48744 (1-748)

QY 1766 CGGTGCTGCTAGTGTTCCTCCGACGACCTTACTGTCTGGATGGTTCTCATCT 1825
|||||
Db 521 AlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySerAlaCys 540
|||||

RESULT 6
G01856
semaphorin V - human
C:Species: Homo sapiens (man)
C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 24-Sep-1999
C:Accession: G01856
R:Sekido, Y.
submitted to the EMBL Data Library, June 1995
A:Reference number: G08634
A:Accession: G01856
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-749 <SEK>
A:Cross-references: EMBL:U28369; NID:9974283; PIDN:AAD09138.1; PID:g974284
C:Superfamily: semaphorin

Alignment Scores:
Pred. No.: 9.73e-07 Length: 749
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x G01856 (1-749)

QY 1028 AGAAGTCTGGTGAATAAATGACACATTCCTCAAGCTCGTCTG 1072
|||||
Db 277 ArgSerLeuValAsnLysTrpThrThrPheLeuLysAlaArgLeu 291
|||||

RESULT 7
I48748
semaphorin E - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C:Accession: I48748
R:Busch, A.W.; Adams, R.H.; Betz, H.
Neuron 14, 941-948, 1995
A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates
A:Reference number: I48744; MUID:95267431; PMID:7748561
A:Accession: I48748
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-751 <RES>
A:Cross-references: EMBL:X85994; NID:9854331; PIDN:CAA59986.1; PID:9854332
C:Genetics:
A:Gene: semaE
C:Superfamily: semaphorin

Alignment Scores:
Pred. No.: 9.73e-07 Length: 751
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x I48748 (1-751)

QY 1028 AGAAGTCTGGTGAATAAATGACACATTCCTCAAGCTCGTCTG 1072
|||||
Db 274 ArgSerLeuValAsnLysTrpThrThrPheLeuLysAlaArgLeu 288
|||||

RESULT 8
G02173
semaphorin III family homolog - human
C:Species: Homo sapiens (man)
C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 24-Sep-1999
C:Accession: G02173

R.Naylor, S.
submitted to the EMBL Data Library, October 1995
A;Reference number: G09275
A;Accession: G02173
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: mRNA
A;Residues: 1-753 <NAY>
A;Cross-references: EMBL:U38276; NID:g1061350; PIDN:AAB18276.1; PID:g1061351
C;Superfamily: semaphorin

Alignment Scores:
Pred. No.: 0.00014 Length: 753
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x G02173 (1-753)
QY 1778 TGTTCCTCGCCGAGACCTTACTGTGCTGGGATGGT 1816
|||||
Db 526 CysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 538
|||||

RESULT 9
T03102
semaphorin homolog A3 - alcelaphine herpesvirus 1
C;Species: alcelaphine herpesvirus 1
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 08-Oct-1999
A;Accession: T03102
R;Jensner, A.; Pflanz, R.; Fleckenstein, B.
J. Virol. 71, 6517-6525, 1997
A;Title: Primary structure of the alcelaphine herpesvirus 1 genome.
A;Reference number: Z14840; MUID:97404659; PMID:9261371
A;Accession: T03102
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-653 <ENS>
A;Cross-references: EMBL:AF005370; NID:g2337967; PIDN:AAC58054.1; PID:g2337970

Alignment Scores:
Pred. No.: 0.244 Length: 653
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x T03102 (1-653)

QY 1043 AAATGGACACATTCCTCAAGCTGCTGTG 1072
|||||
Db 302 LysTrpThrPheLeuLysAlaArgLeu 311
|||||

RESULT 10
S66498
M-sema F protein precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 24-Nov-1999
A;Accession: S66498
R;Inagaki, S.; Furuyama, T.; Iwahaashi, Y.
FEBS Lett. 370, 269-272, 1995
A;Title: Identification of a member of mouse semaphorin family.
A;Reference number: S66498; MUID:95385809; PMID:7656991
A;Accession: S66498
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-834 <INA>
A;Cross-references: EMBL:S79463; NID:g1110598; PIDN:AAB35184.1; PID:g1110599
C;Superfamily: semaphorin
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-834/Product: M-sema F protein #status predicted <MAT>

Alignment Scores:
Pred. No.: 0.237 Length: 834
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x S66498 (1-834)
QY 1043 AAATGGACACATTCCTCAAGCTGCTGTG 1072
|||||
Db 274 LysTrpThrPheLeuLysAlaArgLeu 283
|||||

RESULT 11
T11017
MLCL536.34 protein - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 23-Mar-2001
A;Accession: T11017; S72779
R;Parthill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, September 1997
A;Reference number: Z16918
A;Accession: T11017
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-311 <PAR>
A;Cross-references: EMBL:Z99125; NID:g2398683; PIDN:CAB16178.1; PID:g2398713
R;Smith, D.R.; Robison, K.
submitted to the EMBL Data Library, November 1993
A;Description: Mycobacterium leprae cosmid B1496.
A;Reference number: S72695
A;Accession: S72779
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 'M',90-311 <SMI>
A;Cross-references: EMBL:U00013; NID:g466868; PIDN:AAA17144.1; PID:g466899
C;Genetics:
A;Gene: MLCL536.34
A;Start codon: GTG

Alignment Scores:
Pred. No.: 3.17 Length: 311
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x T11017 (1-311)

QY 516 GCCCACTGCATTCATCTCTCTGTG 490
|||||
Db 225 AlaHisLeuHisSerSerLeuLeuVal 233
|||||

RESULT 12
CIHUS
complement subcomponent C 1S8AR.GIF (EC 3.4.21.42) precursor [validated] - human
N;Alternate names: C1 esterase precursor
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 03-Jun-2002
A;Accession: A40496; A27381; S00224; S05634; A05140; A25396; A38407; B37820
R;Kusumoto, H.; Hirose, S.; Sallier, J.P.; Hagen, F.S.; Kurachi, K.
Proc. Natl. Acad. Sci. U.S.A. 85, 7307-7311, 1988
A;Title: Human genes for complement components C1r and C1s in a close tail-to-tail arrangement
A;Reference number: A40496; MUID:89017187; PMID:2459702
A;Accession: A40496
A;Molecule type: mRNA
A;Residues: 1-688 <KUS>
A;Cross-references: GB:J04080; NID:g179645; PIDN:AAA51852.1; PID:g179646
R;Tosi, M.; Duponchel, C.; Meo, T.; Jullier, C.
Biochemistry 26, 8516-8524, 1987
A;Title: Complete cDNA sequence of human complement C1s and close physical linkage of th

F;16-437/Product: complement subcomponent C1s chain A (heavy chain) #status experimental
F;135-171/Domain: EGF homology <EGF>
F;175-287/Domain: C1r/C1s repeat homology <C1R2>
F;294-354/Domain: complement factor H repeat homology <PH1>
F;359-421/Domain: complement factor H repeat homology <PH2>
F;438-688/Product: complement subcomponent C1s chain B (light chain) #status experimental
F;438-675/Domain: trypsin homology <TRY>
F;65-83,135-147,143-156,158-171,175-202,234-251,294-341,321-354,359-403,386-421,425-549
F;149/Modified site: erythro-beta-hydroxyasparagine (Asn) (partial) #status experimental
F;174,406/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;437-438/Cleavage site: Arg-Ile (complement subcomponent C1r) #status experimental
F;475,529,632/Active site: His, Asp, Ser #status predicted

Alignment Scores:

Pred. No.:	2.9	Length:	688
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.03%	Indels:	0
DB:	1	Gaps:	0

US-09-774-490-1 (1-2709) x C1HUS (1-688)

QY 151 TTGTGCGCCGACGAGAGTTCAACAACAT 125
|||||
DB 82 LeuCysGlyGlnArgSerSerAsnAsn 90
|||||

RESULT 13
I48746
semaphorin C - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 03-Nov-2000
C;Accession: I48746
R;Puschel, A.W.; Adams, R.H.; Betz, H.
Neuron 14, 941-948, 1995
A;Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates
A;Reference number: I48744; MUID:95267431; PMID:7748561
A;Accession: I48746
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-782 <RES>
A;Cross-references: EMBL:X85992; NID:g854327; PIDN:CAA59984.1; PID:g854328
C;Genetics:
A;Gene: semC
C;Superfamily: semaphorin

Alignment Scores:

Pred. No.:	2.86	Length:	782
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.01%	Indels:	0
DB:	2	Gaps:	0

US-09-774-490-1 (1-2709) x I48746 (1-782)

QY 1784 CTCGCGCCGACCCCTACTGCTTGG 1810
|||||
DB 483 LeuAlaArgAspProtyrCysAlaTrp 491
|||||

RESULT 14
S25327
cytoskeleton assembly control protein SLAI - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YBL007c; protein YBL0321
C;Species: Saccharomyces cerevisiae
C;Date: 23-Apr-1993 #sequence_revision 23-Apr-1993 #text_change 29-Oct-1999
C;Accession: S25327; A40673; S45740; S36355; S42277
R;Delaveau, T.; Jacq, C.; Perea, J.
Yeast 8, 761-768, 1992
A;Title: Sequence of a 12.7 kb segment of yeast chromosome II identifies a PDR-like gene
A;Reference number: S25326; MUID:93070613; PMID:1441753
A;Accession: S25327
A;Molecule type: DNA

A;Residues: 1-1244
A;Cross-references: EMBL:S47695; NID:G259049; PIDN:AAB23985.1; PID:G259051
R;Holtzman, D.A.; Yang, S.; Drubin, D.G.
J. Cell Biol. 122, 635-644, 1993
A;Title: Synthetic-lethal interactions identify two novel genes, SLA1 and SLA2, that control cell growth.
A;Reference number: A40673; MUID:93328765; PMID:8335689
A;Accession: A40673
A;Molecule type: DNA
A;Residues: 1-1244 <HOL>
A;Cross-references: GB:222810; NID:G311410; PIDN:CAA80463.1; PID:G311411
R;Delaveau, T.; Jacq, C.; Perea, J.
Submitted to the Protein Sequence Database, August 1994
A;Reference number: S45736
A;Accession: S45740
A;Molecule type: DNA
A;Residues: 1-1244 <DE2>
A;Cross-references: EMBL:Z35768; NID:G535989; PIDN:CAA84826.1; PID:G535990; MIPS:YBL0070
C;Comment: This protein is involved in formation of cortical cytoskeleton.
C;Genetics:
A;Gene: SGD:SLA1
A;Cross-references: SGD:S000103; MIPS:YBL007C
A;Map position: 2L
C;Superfamily: SH3 homology
C;Keywords: Cytoskeleton
F;10-65/Domain: SH3 homology <SH31>
F;76-127/Domain: SH3 homology <SH32>
F;360-410/Domain: SH3 homology <SH33>

Alignment Scores:
Pred. No.: 2.72 Length: 1244
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S25327 (1-1244)

QY 380 CTTTGGATGAGGACGGAGTAGGCTG 406
|||||
DB 607 LeuLeuAepGluGluArgSerArgLeu 615

RESULT 15
S74173
ryanodine receptor 3 - Mustela sp.
C;Species: Mustela sp.
C;Date: 23-Apr-1998 #sequence_revision 01-May-1998 #text_change 18-Jun-1999
C;Accession: S74173
R;Marziali, G.; Rossi, D.; Giannini, G.; Charlesworth, A.; Sorrentino, V.
FEBS Lett. 394, 76-82, 1996
A;Title: cDNA cloning reveals a tissue specific expression of alternatively spliced transcripts.
A;Reference number: S74173; MUID:97074208; PMID:8925932
A;Accession: S74173
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-4859 <MAR>
A;Superfamily: ryanodine receptor; transcription initiation factor sigma region 1, homologous
C;Keywords: alternative splicing; calcium channel; glycoprotein; phosphoprotein; transmembrane
F;697-702/Region: adenine nucleotide binding #status predicted
F;849-934/Domain: repeat #status predicted <RPT1>
F;963-1048/Domain: repeat #status predicted <RPT2>
F;1134-1139/Region: adenine nucleotide binding #status predicted
F;1224-2239/Region: adenine nucleotide binding #status predicted
F;2521-2526/Region: adenine nucleotide binding #status predicted
F;2536-2681/Domain: repeat #status predicted <RPT3>
F;2714-2792/Domain: repeat #status predicted <RPT4>
F;3834-3854/Domain: transmembrane #status predicted
F;3862-3880/Domain: transmembrane #status predicted <TM1>
F;3914-3932/Domain: calcium binding #status predicted <CA1>
F;4119-4142/Domain: transmembrane #status predicted <TM2>
F;4184-4204/Domain: transmembrane #status predicted <TM3>
F;4399-4420/Domain: transmembrane #status predicted <TM4>
F;4399-4420/Domain: transmembrane #status predicted <TM5>

F;4471-4494/Domain: transmembrane #status predicted <TM6>
F;4611-4642/Domain: transmembrane #status predicted <TM7>
F;4659-4678/Domain: transmembrane #status predicted <TM8>
F;4701-4720/Domain: transmembrane #status predicted <TM9>
F;4736-4759/Domain: transmembrane #status predicted <TM10>
F;130,290,1243/Binding site: phosphate (Thr) (covalent) (by cAMP- and calmodulin-dependent kinase)
F;2706/Binding site: phosphate (Ser) (covalent) (by calmodulin-dependent kinase) #status predicted
F;4686/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 2.34 Length: 4859
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S74173 (1-4859)

QY 67 CGTGGTCGACACCCCTCGACGTGGCA 41
|||||
DB 2132 ArgGlySerThrProLeuAspValAla 2140

RESULT 16
B54161
ryanodine-binding protein beta form - bullfrog
C;Species: Rana catesbeiana (bullfrog)
C;Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 20-Jun-2000
C;Accession: B54161
R;Oyamada, H.; Murayama, T.; Takagi, T.; Iino, M.; Iwabe, N.; Miyata, T.; Ogawa, Y.; Endo, J.
J. Biol. Chem. 269, 17206-17214, 1994
A;Title: Primary structure and distribution of ryanodine-binding protein isoforms of the A;Reference number: A54161; MUID:94274714; PMID:8006029
A;Accession: B54161
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-4868 <OYA>
A;Cross-references: GB:D21071; NID:G1856973; PIDN:BA04647.1; PID:G538247
C;Superfamily: ryanodine receptor; transcription initiation factor sigma region 1, homologous

Alignment Scores:
Pred. No.: 2.34 Length: 4868
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x B54161 (1-4868)

QY 67 CGTGGTCGACACCCCTCGACGTGGCA 41
|||||
DB 2136 ArgGlySerThrProLeuAspValAla 2144

RESULT 17
S66572
ryanodine receptor type 3 - chicken
C;Species: Gallus gallus (chicken)
C;Date: 14-Feb-1997 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C;Accession: S66572
R;Ottini, L.; Marziali, G.; Conti, A.; Charlesworth, A.; Sorrentino, V.
Biochem. J. 315, 207-216, 1996
A;Title: Alpha and beta isoforms of ryanodine receptor from chicken skeletal muscle are A;Reference number: S66572; MUID:96207583; PMID:8670108
A;Accession: S66572
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-4869 <OTT>
A;Cross-references: EMBL:X95267; NID:G1212911; PIDN:CAA64563.1; PID:G1212912
C;Superfamily: ryanodine receptor; transcription initiation factor sigma region 1, homologous

Alignment Scores:
Pred. No.: 2.34 Length: 4869

Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S66572 (1-4869)

QY 67 CQTGGTGCACACCCCTCGACGTGGCA 41
Db 2135 ArgGlySerThrProLeuAspValAla 2143

RESULT 18
S27272
ryanodine receptor, brain - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 11-Jun-1999
C:Accession: S27272
R:Hakamata, Y.; Nakai, J.; Takeshima, H.; Imoto, K.
FBS Lett. 312, 229-235, 1992
A:Title: Primary structure and distribution of a novel ryanodine receptor/calcium release
A:Reference number: S27272; MUID:93050200; PMID:1330694
A:Accession: S27272
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-4872 <HAK>
A:Cross-references: EMBL:X68650
C:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolo

Alignment Scores:
Pred. No.: 2.34 Length: 4872
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S27272 (1-4872)

QY 67 CQTGGTGCACACCCCTCGACGTGGCA 41
Db 2136 ArgGlySerThrProLeuAspValAla 2144

RESULT 19
S27269
ryanodine receptor isoform 2, cardiac muscle - human
N:Alternate names: calcium release channel protein isoform 2
C:Species: Homo sapiens (man)
C:Date: 14-Apr-1998 #sequence_revision 24-Apr-1998 #text_change 20-Jun-2000
C:Accession: S72269
R:Tunwell, R.E.A.; Wickenden, C.; Bertrand, B.M.A.; Shevchenko, V.I.; Walsh, M.B.; Allen
Biochem. J. 318, 477-487, 1996
A:Title: The human cardiac muscle ryanodine receptor-calcium release channel: identifica
A:Reference number: S72269; MUID:96404895; PMID:8809036
A:Accession: S72269
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-4967 <TUN>
A:Cross-references: EMBL:X98330; NID:g1526977; PIDN:CA66975.1; PID:g1526978
A:Experimental source: cardiac muscle; sarcoplasmic reticulum
C:Genetics:
A:Gene: RYR-2
C:Complex: homotetramer
C:Function:
A:Description: mediates rapid release of Ca2+ ions from intracellular stores during card
C:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolo
C:Keywords: alternative splicing; calcium channel; cardiac muscle; glycoprotein; heart;
F:710-715/Region: adenine nucleotide binding #status predicted
F:1148-1153/Region: adenine nucleotide binding #status predicted
F:2338-2342/Region: adenine nucleotide binding #status predicted
F:2626-2631/Region: adenine nucleotide binding #status predicted
F:4030-4048/Domain: calcium binding #status predicted <CAL>
F:4499-4519/Domain: transmembrane #status predicted <TM01>

F:4572-4593/Domain: transmembrane #status predicted <TM02>
F:4719-4744/Domain: transmembrane #status predicted <TM03>
F:4767-4789/Domain: transmembrane #status predicted <TM04>
F:4809-4829/Domain: transmembrane #status predicted <TM05>
F:4844-4867/Domain: transmembrane #status predicted <TM06>
F:140-301/Binding site: phosphate (Thr) (covalent) (by cAMP- and calmodulin-dependent kin
F:2808/Binding site: phosphate (Ser) (covalent) (by calmodulin-dependent kinase) #status
F:4794/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.: 2.34 Length: 4967
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S72269 (1-4967)

QY 67 CQTGGTGCACACCCCTCGACGTGGCA 41
Db 2235 ArgGlySerThrProLeuAspValAla 2243

RESULT 20
A37113
ryanodine receptor, cardiac muscle - rabbit
N:Alternate names: ryanodine receptor 2
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 15-Feb-1991 #sequence_revision 15-Feb-1991 #text_change 20-Jun-2000
C:Accession: A37113; S74212
J:Otsu, K.; Willard, H.F.; Khanna, V.K.; Zorzato, F.; Green, N.M.; MacLennan, D.H.
J. Biol. Chem. 265, 13472-13483, 1990
A:Title: Molecular cloning of cDNA encoding the Ca(2+) release channel (ryanodine recept
A:Reference number: A37113; MUID:90337947; PMID:2380170
A:Accession: A37113
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-4969 <OTS>
A:Cross-references: GB:M59743; GB:J05564; NID:g164831; PIDN:AAA31179.1; PID:g164832
R:Nishida, K.; Otsu, K.; Hori, M.; Kuzuya, T.; Tada, M.
Eur. J. Biochem. 240, 408-415, 1996
A:Title: Cloning and characterization of the 5'-upstream regulatory region of the Ca2+-r
A:Reference number: S74212; MUID:96439071; PMID:8841406
A:Accession: S74212
A:Molecule type: DNA
A:Residues: 1-16 <NIS>
A:Cross-references: EMBL:X99486
A:Experimental source: strain New Zealand White
C:Genetics:
A:Gene: RYR2
C:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolo
C:Keywords: cardiac muscle; heart; phosphoprotein; transmembrane protein

Alignment Scores:
Pred. No.: 2.34 Length: 4969
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x A37113 (1-4969)

QY 67 CQTGGTGCACACCCCTCGACGTGGCA 41
Db 2235 ArgGlySerThrProLeuAspValAla 2243

RESULT 21
A54161
ryanodine-binding protein alpha form - bullfrog
C:Species: Rana catesbeiana (bullfrog)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C:Accession: A54161

R.; Oyama, H.; Murayama, T.; Takagi, T.; Iino, M.; Iwabe, N.; Miyata, T.; Ogawa, Y.; Endo, J. Biol. Chem. 269, 17206-17214, 1994
A:Title: Primary structure and distribution of ryanodine-binding protein isoforms of the A:Reference number: A54161; MUID:94274714; PMID:8006029
A:Accession: A54161
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-5037 <OVA>
A:Cross-references: GB:D21070; NID:g1856972; PIDN:BAA04646.1; PID:g538246
C:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolog

Alignment Scores:
Pred. No.: 2.33 Length: 5037
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x A54161 (1-5037)

QY 67 CGTGGTCGACACCCCTCGAGTGGCA 41
|||||
Db 2253 ArgGlySerThrProLeuAspValAla 2261
|||||

RESULT 22
S40450
ryanodine receptor/calcium release channel - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 11-Jun-1999
C:Accession: S40450
R:Takeshima, H.; Mishi, M.; Iwabe, N.; Miyata, T.; Hosoya, T.; Masai, I.; Hotta, Y.
FEBS Lett. 337, 81-87, 1994
A:Title: Isolation and characterization of a gene for a ryanodine receptor/calcium release channel
A:Reference number: S40450; MUID:94102409; PMID:8276118
A:Accession: S40450
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-5126 <TAK>
C:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolog

Alignment Scores:
Pred. No.: 2.33 Length: 5126
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S40450 (1-5126)

QY 67 CGTGGTCGACACCCCTCGAGTGGCA 41
|||||
Db 2355 ArgGlySerThrProLeuAspValAla 2363
|||||

RESULT 23
S02763
pnd protein - Escherichia coli plasmid R16
C:Species: Escherichia coli
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S02763
R:Sakikawa, T.; Akimoto, S.; Ohnishi, Y.
Biochim. Biophys. Acta 1007, 158-166, 1989
A:Title: The pnd gene in E. coli plasmid R16: nucleotide sequence and gene expression level
A:Reference number: S02763; MUID:89150247; PMID:2465777
A:Accession: S02763
A:Molecule type: DNA
A:Residues: 1-50 <SAK>
A:Cross-references: EMBL:X12833; NID:942453; PIDN:CAA31320.1; PID:g42454
C:Genetics:
A:Gene: pnd
A:Genome: plasmid
C:Superfamily: flmA protein

Alignment Scores:
Pred. No.: 46.4 Length: 50
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x S02763 (1-50)

QY 1374 CAAGGACCTTCCTGATGATGTTA 1397
|||||
Db 3 GlnArgThrPheLeuMetMetLeu 10
|||||

RESULT 24
JU0071
pnd protein - Escherichia coli plasmid R483
C:Species: Escherichia coli
C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 20-Jun-2000
C:Accession: JU0071
R:Ono, K.; Akimoto, S.; Ohnishi, Y.
Microbiol. Immunol. 31, 1071-1083, 1987
A:Title: Nucleotide sequence of the pnd gene in plasmid R483 and role of the pnd gene product in the degradation of stable RNA in growing Escherichia coli
A:Reference number: JU0071; MUID:89174457; PMID:3328043
A:Accession: JU0071
A:Molecule type: DNA
A:Residues: 1-50 <ONO>
A:Cross-references: GB:D00364; NID:g216610; PIDN:BAA00270.1; PID:g216612
C:Comment: The protein increases the degradation of stable RNA in growing Escherichia coli
C:Genetics:
A:Genome: plasmid
C:Superfamily: flmA protein

Alignment Scores:
Pred. No.: 46.4 Length: 50
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x JU0071 (1-50)

QY 1374 CAAGGACCTTCCTGATGATGTTA 1397
|||||
Db 3 GlnArgThrPheLeuMetMetLeu 10
|||||

RESULT 25
B38256
tryptophan biosynthetic enzyme transcription regulatory protein mtrB [validated] - Bacillus subtilis
N:Alternate names: trp RNA-binding attenuation protein; tryptophan operon RNA-binding protein
C:Species: Bacillus subtilis
C:Date: 14-Jun-1991 #sequence_revision 14-Jun-1991 #text_change 15-Sep-2000
C:Accession: B38256; H69661
R:Gollnick, P.; Ishino, S.; Kuroda, M.I.; Henner, D.J.; Yanofsky, C.
Proc. Natl. Acad. Sci. U.S.A. 87, 8726-8730, 1990
A:Title: The mtr locus is a two-gene operon required for transcription attenuation in the tryptophan operon
A:Reference number: B38256; MUID:91062353; PMID:2123343
A:Accession: B38256
A:Molecule type: DNA
A:Residues: 1-75 <GOL>
A:Cross-references: GB:M37320; NID:g143230; PIDN:AAA22616.1; PID:g143234
R:Antson, A.A.; Dodson, E.J.; Gollnick, P.
Submitted to the Brookhaven Protein Data Bank, February 1995
A:Reference number: A66916; PDB:1WAP
A:Contents: annotation; X-ray crystallography, 1.80 angstroms
R:Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrarini, E.; Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Iech, J.;

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y.M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Saco, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, Akeuchi, M.; Tanakaka, T.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: H69661
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-75 <XUN>
A:Cross-references: GB:Z99115; GB:AL009126; NID:92634478; PIDN:CAB14193.1; PID:g2634495
A:Experimental source: strain 168
C:Comment: This protein is required for transcription termination in the trp leader region.
C:Genetics:
A:Gene: mtrB
C:Keywords: RNA binding; transcription termination

Alignment Scores:
Pred. No.: 44.4 Length: 75
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x B38256 (1-75)
QY 2034 GCGGAAATGAAGAGCGCAAGAGAG 2057
Db 68 GlyGluMetLysSerGluLysLys 75

RESULT 26
C71047
Hypothetical protein PH1667 - *Pyrococcus horikoshii*
C:Species: *Pyrococcus horikoshii*
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C:Accession: C71047
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, D. Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon.
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: C71047
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-101 <KAW>
A:Cross-references: GB:AP000006; NID:93236133; PIDN:BMA30779.1; PID:g3258096
A:Experimental source: strain OT3
A:Note: This accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH1667
C:Superfamily: *Pyrococcus horikoshii* hypothetical protein PH1667

Alignment Scores:
Pred. No.: 43 Length: 101
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x C71047 (1-101)
QY 67 CGTGGGTGACACCCCTCGACGTG 44
Db 24 ArgGlySerThrProLeuAspVal 31

RESULT 27
AG0244

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y.M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Saco, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, Akeuchi, M.; Tanakaka, T.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: H69661
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-75 <XUN>
A:Cross-references: GB:Z99115; GB:AL009126; NID:92634478; PIDN:CAB14193.1; PID:g2634495
A:Experimental source: strain 168
C:Comment: This protein is required for transcription termination in the trp leader region.
C:Genetics:
A:Gene: mtrB
C:Keywords: RNA binding; transcription termination

Alignment Scores:
Pred. No.: 44.4 Length: 75
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x B38256 (1-75)
QY 2034 GCGGAAATGAAGAGCGCAAGAGAG 2057
Db 68 GlyGluMetLysSerGluLysLys 75

RESULT 26
C71047
Hypothetical protein PH1667 - *Pyrococcus horikoshii*
C:Species: *Pyrococcus horikoshii*
C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C:Accession: C71047
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekin, M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi, D. Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic archaeon.
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: C71047
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-101 <KAW>
A:Cross-references: GB:AP000006; NID:93236133; PIDN:BMA30779.1; PID:g3258096
A:Experimental source: strain OT3
A:Note: This accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH1667
C:Superfamily: *Pyrococcus horikoshii* hypothetical protein PH1667

Alignment Scores:
Pred. No.: 43 Length: 101
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x C71047 (1-101)
QY 67 CGTGGGTGACACCCCTCGACGTG 44
Db 24 ArgGlySerThrProLeuAspVal 31

RESULT 27
AG0244

hypothetical protein YPO2006 [imported] - *Yersinia pestis* (strain CO92)
C:Species: *Yersinia pestis*
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C:Accession: AG0244
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; et al. M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, Nature 413, 523-527, 2001
A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AG0244
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-178 <KUR>
A:Cross-references: GB:AL590842; PIDN:CAC90819.1; PID:g15980020; GSPDB:GN00175
C:Genetics:
A:Gene: YPO2006

Alignment Scores:
Pred. No.: 40.4 Length: 178
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x AG0244 (1-178)
QY 2517 GCTCAGACATCTCTGGTGCCCTC 2494
Db 138 AlaGlnThrLeuLeuGlyAlaLeu 145

RESULT 28
B83932
butyrate-acetoacetate CoA-transferase B82258 [imported] - *Bacillus halodurans* (strain C-1)
C:Species: *Bacillus halodurans*
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: B83932
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira, Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: B83932
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-223 <STO>
A:Cross-references: GB:AP001514; GB:BA000004; NID:g10174613; PIDN:BA05977.1; GSPDB:GN00175
A:Experimental source: strain C-125
C:Genetics:
A:Gene: B82258
C:Superfamily: 3-oxoadipate CoA-transferase beta chain; 3-oxoadipate CoA-transferase beta chain

Alignment Scores:
Pred. No.: 39.4 Length: 223
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x B83932 (1-223)
QY 2120 GTTGTAGACTAGTAGCAGAGGC 2097
Db 36 ValValAspTyrValAlaGluGly 43

RESULT 29
D84868
probable endochitinase [imported] - *Arabidopsis thaliana*
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: D84868
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A>Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: D84868

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-283 <STO>

A:Cross-references: GB:AE002093; NID:G2281108; PIDN:AAB64044.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2G43620

A:Map position: 2

C:Superfamily: lectin-related plant chitinase; hevvin chitin-binding domain homology; pl

Alignment Scores:

Pred. No.:	38.4	Length:	283
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.89%	Indels:	0
DB:	2	Gaps:	0

US-09-774-490-1 (1-2709) x D84868 (1-283)

QY 2319 TCACGCTCATCAACCAACCCCAATC 2342

Db 64 CysSerSerSerThrThrProile 71

RESULT 30

GB7498

hypothetical protein CC2012 [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: GB7498

R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J. B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A>Title: Complete Genome Sequence of *Caulobacter crescentus*.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: GB7498

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-309 <STO>

A:Cross-references: GB:AE005673; NID:G13423483; PIDN:AAK23987.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC2012

Alignment Scores:

Pred. No.:	38	Length:	309
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.92%	Indels:	0
DB:	2	Gaps:	0

US-09-774-490-1 (1-2709) x GB7498 (1-309)

QY 1074 ATCAGACGACTTTCAGGAATGTT 1051

Db 38 IleArgArgAlaLeuArgAsnVal 45

RESULT 31

H97346

oligopeptide ABC transporter, ATPase component CAC3641 [imported] - *Clostridium acetobu*

C:Species: *Clostridium acetobutylicum*

C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001

C:Accession: H97346

R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing *Bacterium Cl*

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: H97346

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-326 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK81563.1; PID:G15026741; GSPDB:GN00168

A:Experimental source: *Clostridium acetobutylicum* ATCC824

C:Genetics:

A:Gene: CAC3641

Alignment Scores:

Pred. No.:	37.8	Length:	326
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.89%	Indels:	0
DB:	2	Gaps:	0

US-09-774-490-1 (1-2709) x H97346 (1-326)

QY 296 AAATTATCTCTACAAAGAAATGTTG 319

Db 84 LysLeuSerTyrLysGluMetLeu 91

RESULT 32

C97346

oligopeptide ABC transporter, ATPase component CAC3635 [imported] - *Clostridium acetobu*

C:Species: *Clostridium acetobutylicum*

C>Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 17-May-2002

C:Accession: C97346

R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee, J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.

J. Bacteriol. 183, 4823-4838, 2001

A>Title: Genome Sequence and Comparative Analysis of the Solvent-Producing *Bacterium Cl*

A:Reference number: A96900; MUID:21359325; PMID:21359325

A:Accession: C97346

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-350 <KUR>

A:Cross-references: GB:AE001437; PIDN:AAK81558.1; PID:G15026736; GSPDB:GN00168

A:Experimental source: *Clostridium acetobutylicum* ATCC824

C:Genetics:

A:Gene: CAC3635

C:Superfamily: inner membrane protein malK; ATP-binding cassette homology

Alignment Scores:

Pred. No.:	37.5	Length:	350
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	0.89%	Indels:	0
DB:	2	Gaps:	0

US-09-774-490-1 (1-2709) x C97346 (1-350)

QY 296 AAATTATCTCTACAAAGAAATGTTG 319

Db 108 LysLeuSerTyrLysGluMetLeu 115

RESULT 33

I39856

spore germination apparatus protein gerBB - *Bacillus subtilis*

C:Species: *Bacillus subtilis*

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000

C:Accession: I39856; A69630

R:Corfe, B.M.; Sammons, R.L.; Smith, D.A.; Mauel, C. Microbiology 140, 471-478, 1994

A>Title: The gerB region of the *Bacillus subtilis* 168 chromosome encodes a homologue of

A:Reference number: I39854; MUID:94282292; PMID:8012571

A:Accession: I39856

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-368 <RES>

A:Cross-references: GB:116960; NID:G289274; PIDN:AAA22467.1; PID:G289276
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
 C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
 A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen
 iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
 Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
 Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
 akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,
 T.; Winters, P.; Wipat, H.F.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:198044033; PMID:9384377

A:Accession: A69630
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-368 <KUN>
 A:Cross-references: GB:Z99122; GB:AL009126; NID:G2636029; PIDN:CAB15598.1; PID:G2636107
 A:Experimental source: strain 168
 A:Note: Germination response to the combination of glucose, fructose, L-asparagine, and
 C:Genetics:
 A:Gene: gerB-2
 A:Map position: 314 deg.
 C:Superfamily: spore germination protein

Alignment Scores:
 Pred. No.: 37.3 Length: 368
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x I39856 (1-368)

QY 2481 GTCCTCCCTGTTCTACCTTCCTTA 2458

DB 56 ValLeuLeuPheLeuProPheLeu 63

RESULT 34
 A59273
 Proteinase inhibitor 8 - human
 N:Alternate names: cytoplasmic antiproteinase 2 (CAP-2)
 C:Species: Homo sapiens (man)
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C:Accession: A59273
 R:Spracher, C.A.; Morgenstern, K.A.; Mathewes, S.; Dahlen, J.R.; Schrader, S.K.; Foster,
 J. Biol. Chem. 270, 29854-29861, 1995
 A:Title: Molecular cloning, expression, and partial characterization of two novel member
 A:Reference number: A59273; MUID:96102039; PMID:8530382
 A:Accession: A59273
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-374 <SPR>
 A:Cross-references: GB:L40377; NID:gl160927; PIDN:AC41939.1; PID:gl160927
 A:Experimental source: tissue type placenta; note (vector lambda gt11); gene CAP2
 C:Genetics:
 A:Gene: GDB:P18; CAP2
 A:Cross-references: GDB:599392; OMIM:601697
 A:Map position: 18q21.3-18q21.3
 C:Superfamily: antithrombin III
 C:Keywords: proteinase inhibitor
 F:339/Inhibitory site: Arg (unidentified proteinase) #status predicted

Alignment Scores:
 Pred. No.: 37.2 Length: 374
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0

Query Match: 0.89% Indels: 0
 DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x A59273 (1-374)

QY 1590 TTCCTAAGGAGACTTGGTATGATT 1613

DB 282 PheLeuArgArgLeuGlyMetIle 289

RESULT 35

B30341

G protein-coupled receptor RDC4 - dog
 C:Species: Canis lupus familiaris (dog)
 C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 05-Nov-1999
 C:Accession: B30341; NID:9222
 R:Libert, F.; Parmentier, M.; Lefort, A.; Dinsart, C.; Van Sande, J.; Maenhaut, C.; Sino
 Science 244, 569-572, 1989
 A:Title: Selective amplification and cloning of four new members of the G protein-couple
 A:Reference number: A30341; MUID:89242119; PMID:2541503
 A:Accession: B30341

A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-377 <LIB>
 A:Cross-references: EMBL:X14049; NID:9900; PIDN:CAA32207.1; PID:9901
 R:Libert, F.; Parmentier, M.; Lefort, A.; Dumont, J.E.; Vassart, G.
 Nucleic Acids Res. 18, 1916, 1990
 A:Title: Complete nucleotide sequence of a putative G protein coupled receptor: RDC4.
 A:Reference number: S12822; MUID:90245610; PMID:2159630
 A:Accession: S12822

A:Status: translation not shown

A:Molecule type: mRNA

A:Residues: 1-377 <L12>

A:Cross-references: EMBL:X14049; NID:9900; PIDN:CAA32207.1; PID:9901

C:Genetics:
 A:Gene: RDC4
 C:Superfamily: octopamine receptor type I
 C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein

Alignment Scores:
 Pred. No.: 37.2 Length: 377
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x B30341 (1-377)

QY 1917 ATATACCATGGCCACAGCCCTG 1940

DB 47 IleIleThrMetAlaThrAlaLeu 54

RESULT 36

T09461

Type I site-specific deoxyribonuclease (SC 3.1.21.3) LldI chain Hsds - Lactococcus lacti
 C:Species: Lactococcus lactis subsp. lactis

C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 16-Jul-1999

C:Accession: T09461

R:Deng, Y.M.; Dunn, N.W.

A:Submitted to the EMBL Data Library, November 1997

A:Description: LldI, a type I restriction modification system in lactococcus lactis biov

A:Reference number: 216679

A:Accession: T09461

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-409 <DEN>

A:Cross-references: EMBL:AF034786; NID:G2689697; PID:G2689700

A:Experimental source: biovar. diacetylactis; strain UK19161

C:Genetics:
 A:Gene: hsdS
 A:Genome: plasmid pND861
 C:Keywords: hydrolase

Alignment Scores:
Pred. No.: 36.8 Length: 409
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x T09461 (1-409)

QY 1135 GAATACATCTCTGAGTTTCATCAA 1112

Db 117 GluTyrIleLeuGlnPheIleLeu 124

RESULT 37

T22969
hypothetical protein F59A1.13 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T22969

R:Mortimore, B.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z19644

A:Accession: T22969

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-421 <WIL>

A:Cross-references: EMBL:Z81557; PIDN:CAB04538.1; GSPDB:GN000023; CESP:F59A1.13

A:Experimental source: clone F59A1

C:Genetics:

A:Gene: CESP:F59A1.13

A:Map position: 5

A:Introns: 27/1; 116/1; 245/3; 286/3; 340/3; 381/3

Alignment Scores:
Pred. No.: 36.7 Length: 421
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x T22969 (1-421)

QY 291 CTGGCACATGTTCTTCCCATTC 268

Db 77 LeuGlyThrLeuPhePheProPhe 84

RESULT 38

JS0376
hypothetical 53.0K protein - soybean chlorotic mottle virus

C:Species: soybean chlorotic mottle virus

A:Note: host Glycine max (soybean)

C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 28-Jul-2000

C:Accession: JS0376

R:Haegawa, A.; Verver, J.; Shimada, A.; Saito, M.; Goldbach, R.; van Kammen, A.; Miki,

Nucleic Acids Res. 17, 9993-10013, 1989

A:Title: The complete sequence of soybean chlorotic mottle virus DNA and the identification

A:Reference number: JS0372; MUID:90098857; PMID:2602148

A:Accession: JS0376

A:Molecule type: DNA

A:Residues: 1-463 <HAS>

A:Cross-references: GB:X15828; NID:G58833; PIDN:CAA33831.1; PID:G58840

C:Superfamily: soybean chlorotic mottle virus hypothetical 53.0K protein

Alignment Scores:
Pred. No.: 36.3 Length: 463
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x JS0376 (1-463)

QY 1338 CTTGTCCAGCAAAACATTTGGTG 1361

Db 303 LeuValProAlaLysHisLeuVal 310

RESULT 39

G95379

probable integral membrane transporter Sma1717 [imported] - Sinorhizobium meliloti (strain
C:Species: Sinorhizobium meliloti

C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001

C:Accession: G95379

R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows

; Kalman, S.; Keating, D.H.; Palm, C.; Peck, K.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti

A:Reference number: A95262; MUID:21396509; PMID:11481432

A:Accession: G95379

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-475 <KUR>

A:Cross-references: GB:AB006469; PIDN:AAK65601.1; PID:g14524083; GSPDB:GN00165

A:Experimental source: strain 1021, megaplasmid pSymA

R:Galibert, F.; Finan, I.M.; Long, S.R.; Fuhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: Sma1717

A:Genome: plasmid

C:Superfamily: conserved hypothetical protein HI0125

Alignment Scores:
Pred. No.: 36.2 Length: 475
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x G95379 (1-475)

QY 2478 CTCCTGTTTCTACCTTCTTATTT 2455

Db 370 LeuLeuPheLeuProPheLeuPhe 377

RESULT 40

A64117

serine-type D-Ala-D-Ala carboxypeptidase homolog - Haemophilus influenzae (strain Rd KW2

C:Species: Haemophilus influenzae

C:Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 17-Mar-2000

C:Accession: A64117

R:Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage, J.A.

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman, J.

; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.B.; Geoghagen, N.S.M.

Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,

A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.

A:Reference number: A64000; MUID:95350630; PMID:7542800

A:Accession: A64117

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-479 <TIGR>

A:Cross-references: GB:U32812; GB:I42023; NID:g1574784; PIDN:AAC22975.1; PID:g1574789; T

C:Superfamily: D-alanyl-D-alanine carboxypeptidase

Alignment Scores:
Pred. No.: 36.2 Length: 479

Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x A64117 (1-479)

QY 1951 GATTCTCTCTCAGGCTGCGCC 1928

Db 316 AspSerLeupheArgAlaValAla 323

RESULT 41

AB1274

B. subtilis negative regulator of PtsZ ring formation (EzrA) homolog lmo1594 [imported]

C:Species: Listeria monocytogenes

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C:Accession: AB1274

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AB1274

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-571 <GLA>

A:Cross-references: GB:NC_003210; PIDN:CAC99672.1; PID:g16411023; GSPDB:GN00177

A:Experimental source: strain EGD-e

C:Genetics:

A:Gene: lmo1594

Alignment Scores:

Pred. No.: 35.5 Length: 571
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x AB1274 (1-571)

QY 1610 GATTCTCTCTCAGGCTGCGAA 1633

Db 82 AspLeuGluValLeuLeuGlu 89

RESULT 42

AB1637

B. subtilis negative regulator of PtsZ ring formation (EzrA) homolog lin1636 [imported]

C:Species: Listeria innocua

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C:Accession: AC1637

R:Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AC1637

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-571 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC96867.1; PID:g16414123; GSPDB:GN00178

A:Experimental source: strain Clip11262

C:Genetics:

A:Gene: lin1636

Alignment Scores:

Pred. No.: 35.5 Length: 571
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x AC1637 (1-571)

QY 1610 GATTCTCTCTCAGGCTGCGAA 1633

Db 82 AspLeuGluValLeuLeuGlu 89

RESULT 43

SS1528

D-lactate dehydrogenase (cytochrome) (EC 1.1.2.4) - yeast (Kluyveromyces marxianus var.

C:Species: Kluyveromyces marxianus var. lactis, Candida sphaerica

C:Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 29-Oct-1999

C:Accession: SS1528

R:Lodi, T.; O'Connor, D.; Goffrini, P.; Ferrero, I.

Mol. Gen. Genet. 244, 622-629, 1994

A:Title: Carbon catabolite repression in Kluyveromyces lactis: isolation and characteriz

A:Reference number: SS1528; MUID:95059916; PMID:7969031

A:Accession: SS1528

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-579 <LOD>

A:Cross-references: EMBL:X71628; NID:g602028; PIDN:CAA50635.1; PID:g602029

A:Note: the source is designated as Kluyveromyces lactis

C:Keywords: oxidoreductase

Alignment Scores:

Pred. No.: 35.5 Length: 579
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x SS1528 (1-579)

QY 495 GAAGAGATGAATCACTGGCTG 518

Db 470 GluGluMetAsnAlaSerGlyLeu 477

RESULT 44

T37827

hypothetical protein SPAC17A5.12 - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999

C:Accession: T37827

R:Devlin, K.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z21737

A:Accession: T37827

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-697 <DEV>

A:Cross-references: EMBL:Z98849; PIDN:CAB11512.1; GSPDB:GN00066; SPDB:SPAC17A5.12

A:Experimental source: strain 972h-; cosmid c17A5

C:Genetics:

A:Gene: SPDB:SPAC17A5.12

A:Map position: 1

Alignment Scores:

Pred. No.: 34.7 Length: 697
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x T37827 (1-697)

QY 383 TTGGATGAGGAACGAGTAGGCTG 406
DB 592 LeuAspGluGluArgSerArgLeu 599

RESULT 45

S50130
phosphate acetyltransferase (EC 2.3.1.8) [validated] - Escherichia coli (strain K-12)
N:Alternate names: phosphotransacetylase
C:Species: Escherichia coli
C>Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 20-Jun-2000
C:Accession: S50130
R:Matsuyama, A.; Yamamoto-Otake, H.; Hewitt, J.; MacGillivray, R.T.A.; Nakano, E.
Biochim. Biophys. Acta 1219, 559-562, 1994
A:Title: Nucleotide sequence of the phosphotransacetylase gene of Escherichia coli strain K-12
A:Reference number: S50130; MUID:95002178; PMID:7918659
A:Accession: S50130
A:Molecule type: DNA
A:Residues: 1-713 <MAT>
A:Cross-references: EMBL:D21123; NID:9577281; PIDN:BAA04663.1; PID:G501935
A:Experimental source: strain K-12 1100
A>Note: the authors translated the codon CGT for residue 113 as Glu, GTA for residue 407
C:Note: part of this sequence, including the amino end, was confirmed by protein sequencing
C:Genetics: pta
A:Gene: pta
A:Start codon: GTG
C:Function:
A:Description: EC 2.3.1.8 [validated, MUID:95002178]
C:Superfamily: phosphate acetyltransferase pta
C:Keywords: acyltransferase; coenzyme A
F:2-713/Product: phosphate acetyltransferase #status experimental <MAT>

Alignment Scores:
Pred. No.: 34.7 Length: 713
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S50130 (1-713)

QY 1788 GCGAGCGAAGCTGACGACGCT 1765
DB 260 AlaArgGlnHisSerAlaHisAla 267

RESULT 46

T33751
hypothetical protein R11E3.8 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
C:Accession: T33751
R:Langston, Z.; Wohldmann, P.
submitted to the EMBL Data Library, October 1998
A:Description: The sequence of C. elegans cosmid R11E3.
A:Reference number: Z21397
A:Accession: T33751
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-745 <LAN>
A:Cross-references: EMBL:AF100669; PIDN:AC68993.1; GSPDB:GN00022; CESP:R11E3.8
A:Experimental source: strain Bristol N2; clone R11E3
C:Genetics:
A:Gene: CESP:R11E3.8
A:Map position: 4
A:Introns: 104/2; 209/3; 379/3; 528/2; 576/3; 711/2
C:Superfamily: acylaminoacyl-peptidase

Alignment Scores:
Pred. No.: 34.5 Length: 745
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x T33751 (1-745)

QY 1362 CCACCAATGTTTGGTCGGACAA 1339

DB 447 ProProAnValLeuLeuGlyGln 454

RESULT 47

I48745
semaphorin B - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Nov-1999
C:Accession: I48745
R:Puschel, A.W.; Adams, R.H.; Betz, H.
Neuron 14, 941-948, 1995
A:Title: Murine semaphorin D/collapsin is a member of a diverse gene family and creates a guidance cue for axon growth
A:Reference number: I48744; MUID:95267431; PMID:7748561
A:Accession: I48745
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-760 <RES>
A:Cross-references: EMBL:X85991; NID:9854325; PIDN:CAA59983.1; PID:9854326
C:Genetics:
A:Gene: semB
C:Superfamily: semaphorin

Alignment Scores:
Pred. No.: 34.4 Length: 760
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x I48745 (1-760)

QY 578 ACTCACTGTACGCTGTGGAACG 601

DB 137 ThrHisLeuTyrAlaCysGlyThr 144

RESULT 48

T51787
hypothetical protein F28D10_90 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000
C:Accession: T51787
R:Deiseny, M.; Berger, C.; Cooke, R.; Grellet, F.; Laudie, M.; Mewes, H.W.; Rudd, S.; Le;
submitted to the Protein Sequence Database, August 2000
A:Reference number: Z25454
A:Accession: T51787
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-817
A:Cross-references: EMBL:AL391254
A:Experimental source: cultivar Columbia; BAC clone F28D10
C:Genetics:
A:Map position: 3
A:Introns: 88/3; 441/3
A:Note: F28D10_90

Alignment Scores:
Pred. No.: 34.1 Length: 817
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x T51787 (1-817)

QY 1590 TTCCTAAGGAGACTTGGTATGATT 1613

Db 74 PheLeuArgArgLeuGlyMetIle 81
|||||
RESULT 49
A86404
probable protein ATP-dependent DNA helicase RecQ [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: A86404
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, T.H.; Dewar, K.;
ansen, N.P.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Lin, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A86404
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-941 <STO>
A:Cross-references: GB:AE005172; NID:g1098928; PIDN:AG26068.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
Alignment Scores:
Pred. No.: 33.6 Length: 941
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x A86404 (1-941)
QY 991 TCTACGTCAGTAGCTTTCCAGA 968
Db 98 SerSerValSerSerPheSerArg 105
|||||
RESULT 50
A33626
fibrinogen alpha chain - sea lamprey (fragment)
C:Species: Petromyzon marinus (sea lamprey)
C>Date: 09-Mar-1990 #sequence_revision 19-Feb-1999 #text_change 20-Aug-1999
C:Accession: A33626; A03120
R:Wang, Y.Z.; Patterson, J.; Gray, J.E.; Yu, C.; Cottrell, B.A.; Shimizu, A.; Graham, D.
Biochemistry 28, 9801-9806, 1989
A:Title: Complete sequence of the lamprey fibrinogen alpha-chain.
A:Reference number: A33626; MUID:90122867; PMID:2611265
A:Accession: A33626
A:Molecule type: DNA; mRNA; protein
A:Residues: 1-966 <WAN>
A:Cross-references: GB:M30123; NID:g213197; PIDN:AAA49263.1; PID:g213198
A:Note: residues 1-5 are not translated in Figure 2; parts of this sequence, including t
R:Cottrell, B.A.; Doolittle, R.F.
Biochim. Biophys. Acta 453, 426-438, 1976
A:Title: Amino acid sequences of lamprey fibrinopeptides A and B and characterization of
A:Reference number: A03120; MUID:77065679; PMID:999898
A:Accession: A03120
A:Molecule type: protein
A:Residues: 6-11 <COT>
C:Superfamily: fibrinogen alpha chain; fibrinogen disulfide ring homology
C:Keywords: blood coagulation; glycoprotein; plasma
F:6-966/Product: fibrinogen alpha chain #status experimental <PRE>
F:36-164/Domain: fibrinogen disulfide ring homology <FDR>
F:920/Binding site: carbohydrate (Asn) (covalent) #status predicted
Alignment Scores:
Pred. No.: 33.5 Length: 966
Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x A33626 (1-966)
QY 153 ACAGGAGGAGACTAAGCAGCA 176
Db 949 ThrGlyArgLysThrLysAlaAla 956
|||||
RESULT 51
JC5928
semaphorin F precursor - human
C:Species: Homo sapiens (man)
C>Date: 10-Apr-1998 #sequence_revision 08-May-1998 #text_change 17-Nov-2000
C:Accession: JC5928
R:Simmons, A.D.; Pueschel, A.W.; McPherson, J.D.; Overhauser, J.; Lovett, M.
Biochem. Biophys. Res. Commun. 242, 685-691, 1998
A:Title: Molecular cloning and mapping of human semaphorin F from the Cri-du-chat candid
A:Reference number: JC5928; MUID:98125554; PMID:9464278
A:Accession: JC5928
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1074 <SIM>
A:Cross-references: GB:U52840; NID:g2772583; PIDN:AAC09473.1; PID:g2772584
A:Experimental source: brain
C:Comment: This protein disrupts normal brain development and leads to some of the featu
C:Genetics:
A:Gene: semaf
C:Superfamily: human semaphorin F; thrombospondin type 1 repeat homology
F:1-20/Domain: signal sequence #status predicted <SIG>
F:50-533/Domain: semaphorin #status predicted <SEM>
F:840-896/Domain: thrombospondin type 1 repeat homology <THR3>
F:971-993/Domain: transmembrane #status predicted <TMM>
Alignment Scores:
Pred. No.: 33.1 Length: 1074
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x JC5928 (1-1074)
QY 929 TACTTTTCTTCCTGAAATGCA 952
Db 230 TyrPhePhePheArgGluAsnAla 237
|||||
RESULT 52
F84487
probable ABC transporter [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: F84487
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: F84487
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1146 <STO>
A:Cross-references: GB:AE002093; NID:g5001458; PIDN:AAD37023.1; GSPDB:GN00139
C:Genetics:
A:Gene: Atg07680
A:Map position: 2
Alignment Scores:
Pred. No.: 32.9 Length: 1146

Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x F84487 (1-1146)

QY 1966 CTCTACACATAGATGATCTCTC 1943

Db 693 LeuTyThrIleAspSerIeu 700

RESULT 53

A4652
 hypotheical protein At2g25730 [imported] - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: A84652
 R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.;
 euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
 Nature 402, 761-768, 1999
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
 A:Reference number: A84420; MUID:20083487; PMID:10617197
 A:Accession: A84652
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-2420 <STO>
 A:Cross-references: GB:AB002093; NID:g4874311; PIDN:AA031373.1; GSPDB:GN00139
 C:Genetics:
 A:Gene: At2g25730
 A:Map position: 2

Alignment Scores:

Pred. No.:	Length:	Matches:
Score:	30.3	2420
Percent Similarity:	8.00	8
Best Local Similarity:	100.00%	Conservative: 0
Query Match:	100.00%	Mismatches: 0
DB:	0.89%	Indels: 0
	2	Gaps: 0

US-09-774-490-1 (1-2709) x A84652 (1-2420)

QY 1996 GAAGTCGACAGACGCTGGCTA 2019

Db 1467 GluValAlaGluSerAlaGlyIeu 1474

RESULT 54

A35041
 ryanodine receptor type 1, skeletal muscle - human
 N:Alternate names: calcium release channel protein
 C:Species: Homo sapiens (man)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jan-2000
 C:Accession: A35041; I84622; S66630
 R:Zorzato, F.; Fujii, J.; Otsu, K.; Phillips, M.; Green, N.M.; Lai, F.A.; Meissner, G.;
 J. Biol. Chem. 265, 2244-2256, 1990
 A:Title: Molecular cloning of cDNA encoding human and rabbit forms of the Ca(2+) release
 A:Reference number: A35041; MUID:90130482; PMID:2298749
 A:Accession: A35041
 A>Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-5032 <ZOR>
 A:Cross-references: GB:J05200; NID:g337721; PIDN:AAA60294.1; PID:g337722
 R:Otsu, K.; Phillips, M.S.; Khanna, V.K.; de Leon, S.; MacLennan, D.H.
 Genomics 13, 835-837, 1992
 A:Title: Refinement of diagnostic assays for a probable causal mutation for porcine and
 A:Reference number: A46644; MUID:92347887; PMID:1639409
 A:Accession: I84622
 A>Status: translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 598-722 <RES>
 A:Cross-references: GB:M91455; NID:g337723; PIDN:AAA60295.1; PID:g553643
 R:Lynn, S.; Morgan, J.M.; Lamb, H.K.; Meissner, G.; Gillespie, J.I.

FEBS Lett. 372, 6-12, 1995

A:Title: Isolation and partial cloning of ryanodine-sensitive Ca(2+) release channel pro
 A:Reference number: S66630; MUID:96032536; PMID:7556644

A:Accession: S66630

A:Molecule type: mRNA

A:Residues: 4690-4968 <LYN>

A:Experimental source: myometrial smooth muscle

C:Genetics:

A:Gene: GDB:RYR1

A:Cross-references: GDB:I20359; OMIM:180901

A:Map position: 19q13.1-19q13.1

A:introns: 642/2

C:Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolo
 C:Keywords: calcium channel; homotetramer; phosphoprotein; skeletal muscle; transmembran
 F:1788-1984/Domain: transcription initiation factor sigma region 1 homology <SR1>

Alignment Scores:

Pred. No.:	Length:	Matches:
Score:	28	5032
Percent Similarity:	8.00	8
Best Local Similarity:	100.00%	Conservative: 0
Query Match:	100.00%	Mismatches: 0
DB:	0.92%	Indels: 0
	1	Gaps: 0

US-09-774-490-1 (1-2709) x A35041 (1-5032)

QY 64 GGGTCGACACCCCTCGACGTGGCA 41

Db 2268 GlySerThrProLeuAspValAla 2275

RESULT 55

I46646

ryanodine receptor, skeletal muscle - pig

N:Alternate names: calcium release channel protein

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999

C:Accession: I46646; I46645; S31395; I47133; S26624; A37105; I47212; S18135

R:Fujii, J.; Otsu, K.; Zorzato, F.; De Leon, S.; Khanna, V.K.; Weiler, J.E.; O'Brien, P.

Science 253, 448-451, 1991

A:Title: Identification of a mutation in porcine ryanodine receptor associated with mal

A:Reference number: I46645; MUID:91320118; PMID:1862346

A:Accession: I46646

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-5035 <FUJ>

A:Cross-references: GB:M91452; NID:g164647; PIDN:AAA31119.1; PID:g164648

A:Accession: I46645

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-614,'C',616-5035 <FU2>

A:Cross-references: GB:M91451; NID:g164645; PIDN:AAA31118.1; PID:g164646

R:Leeb, T.; Brem, G.; Brenig, B.

submitted to the EMBL Data Library, November 1992

A:Description: Genomic organization of porcine skeletal muscle ryanodine receptor gene c

A:Reference number: S31395

A:Accession: S31395

A:Molecule type: DNA

A:Residues: 1542-2643 <LEE>

A:Cross-references: EMBL:X69465

R:Leeb, T.; Schmoelzl, S.; Brem, G.; Brenig, B.

Genomics 18, 349-354, 1993

A:Title: Genomic organization of the porcine skeletal muscle ryanodine receptor (RYR1) g

A:Reference number: A48915; MUID:94117003; PMID:8288238

A:Contents: annotation

R:Harbitz, I.; Kristensen, T.; Bosnes, M.; Kran, S.; Davies, W.

Anim. Genet. 23, 395-402, 1992

A:Title: DNA sequence of the skeletal muscle calcium release channel cDNA and verificati

A:Reference number: I47133; MUID:93036581; PMID:1329581

A:Accession: I47133

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 2-2091,'A',2093-3494,'L',3496-4163,'S',4165-4182,'R',4184-4411,'W',4413-4971

A;Cross-references: EMBL:X62880; NID:g1936; PIDN:CAA44674.1; PID:g1937
 R;Harbitz, I.; Kristensen, T.; Kran, T.; Davies, W.
 submitted to the EMBL Data Library, August 1992
 A;Reference number: S26624
 A;Accession: S26624
 A;Molecule type: DNA
 A;Residues: 482-706 <RAW>
 A;Cross-references: EMBL:X68247
 R;Harbitz, I.; Chowdhary, B.; Thomsen, P.D.; Davies, W.; Kaufmann, U.; Kran, S.; Gustava
 Genomics 8, 243-248, 1990
 A;Title: Assignment of the porcine calcium release channel gene, a candidate for the mal
 A;Reference number: A37105; MUID:91065640; PMID:2174405
 A;Accession: A37105
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 4785-4971, 'R', 4973-5035 <HA2>
 A;Cross-references: GB:IM32501; NID:g164428; PIDN:AAA31022.1; PID:g164429
 R;Liedbetter, M.W.; Gruber, J.K.; Louis, C.F.; Mickelson, J.R.
 J. Biol. Chem. 269, 31544-31551, 1994
 A;Title: Tissue distribution of ryanodine receptor isoforms and alleles determined by re
 A;Reference number: A55660; MUID:95081095; PMID:7989322
 A;Accession: I47212
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 516-676 <LED>
 A;Cross-references: EMBL:U15965; NID:g562095; PIDN:AAA60467.1; PID:g562096
 C;Genetics:
 A;Gene: RYR1
 A;Introns: 527/1; 559/1; 598/3; 643/2; 1570/3; 1646/2; 1850/3; 1939/3; 2006/3; 2044/1; 2
 A;Note: the list of introns may be incomplete
 A;Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolo
 C;Keywords: calcium channel
 Alignment Scores:
 Pred. No.: 28 Length: 5035
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x I46646 (1-5035)
 QY 64 GGGTCGACACCCCTCGACGTGGCA 41
 Db 2270 GlySerThrProLeuAspValAla 2277
 RESULT 56
 B35041
 ryanodine receptor, skeletal muscle - rabbit
 N;Alternate names: calcium-release channel protein; junctional channel complex
 C;Species: Oryctolagus cuniculus (domestic rabbit)
 C;Date: 27-Jul-1990 #sequence revision 10-Mar-1994 #text_change 20-Aug-1999
 C;Accession: S04654; B35041; A36181; S53794; S32504
 R;Takeshima, H.; Nishimura, S.; Matsumoto, T.; Ishida, H.; Kangawa, K.; Minamino, N.; Ma
 Nature 339, 439-445, 1989
 A;Title: Primary structure and expression from complementary DNA of skeletal muscle ryan
 A;Reference number: S04654; MUID:89262082; PMID:2725677
 A;Accession: S04654
 A;Molecule type: mRNA
 A;Residues: 1-5037 <TAK>
 A;Cross-references: EMBL:X15750; NID:g1709; PIDN:CAA33762.1; PID:g1710
 A;Note: part of this sequence was confirmed by protein sequencing
 R;Zorzato, F.; Fujii, J.; Otsu, K.; Phillips, M.; Green, N.M.; Lai, F.A.; Meisner, G.;
 J. Biol. Chem. 265, 2244-2256, 1990
 A;Title: Molecular cloning of cDNA encoding human and rabbit forms of the Ca(2+) release
 A;Reference number: A35041; MUID:90130482; PMID:2298749
 A;Accession: B35041
 A;Status: not compared with conceptual translation
 A;Molecule type: mRNA
 A;Residues: 1-66, 'Y', 68-2014, 'D', 2016-3246, 'E', 3248-3480, 3486-4497, 'LE', 4498-4521, 'Q', 45
 , 4748-4758, 'N', 4760-5037 <ZOR>
 R;Marks, A.R.; Tempst, P.; Hwang, K.S.; Taubman, M.B.; Inui, M.; Chadwick, C.; Fleischer

Proc. Natl. Acad. Sci. U.S.A. 86, 8683-8687, 1989
 A;Title: Molecular cloning and characterization of the ryanodine receptor/junctional cha
 A;Reference number: A36181; MUID:90046857; PMID:2813419
 A;Accession: A36181
 A;Status: nucleic acid sequence not shown; not compared with conceptual translation
 A;Molecule type: mRNA; protein
 A;Residues: 'X', 1223-1235, 'XX', 1238-1251; 1334-1348; 1566-1569, 'X', 1571, 'X', 1573; 1597-1604
 -4679, 'X', 4681-4689, 'X', 4691-4693, 'X', 4695, 'X', 4697-4700 <MAR>
 A;Note: the proteolytic fragments sequenced here from the junctional channel complex as
 R;Varsanyi, M.; Meyer, H.E.
 Biol. Chem. Hoppe-Seyler 376, 45-49, 1995
 A;Title: Sarcoplasmic reticular Ca(2+) release channel is phosphorylated at serine 2843
 A;Reference number: S53794; MUID:95336639; PMID:7612188
 A;Accession: S53794
 A;Molecule type: protein
 A;Residues: 2841-2852 <VAR>
 R;Takeshima, H.; Nishimura, S.; Nishi, M.; Ikeda, M.; Sugimoto, T.
 FEBS Lett. 322, 105-110, 1993
 A;Title: A brain-specific transcript from the 3'-terminal region of the skeletal muscle
 A;Reference number: S32504; MUID:93245969; PMID:8097730
 A;Accession: S32504
 A;Molecule type: mRNA
 A;Residues: 4163-5037 <TAW>
 C;Superfamily: ryanodine receptor; transcription initiation factor sigma region 1 homolo
 C;Keywords: calcium channel; glycoprotein; phosphoprotein; skeletal muscle; transmembran
 P;695-704/Region: adenine nucleotide binding
 P;841-954, 955-1068/Region: 114-residue repeats
 P;1344-1359, 1371-1386/Region: 16-residue repeats
 P;1789-1985/Domain: transcription initiation factor sigma region 1 homology <SRI>
 P;2370-2375/Region: adenine nucleotide binding
 P;2725-2844, 2845-2958/Region: 120-residue repeats
 P;4564-4580/Domain: transmembrane #status predicted <TM1>
 P;4641-4664/Domain: transmembrane #status predicted <TM2>
 P;4836-4859/Domain: transmembrane #status predicted <TM3>
 P;4918-4937/Domain: transmembrane #status predicted <TM4>
 P;128, 286/Binding site: phosphate (Thr) (covalent) #status experimental
 P;2843/Binding site: phosphate (Ser) (covalent) #status experimental
 P;4864/Binding site: carboxylate (Asn) (covalent) #status predicted
 Alignment Scores:
 Pred. No.: 28 Length: 5037
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 2 Gaps: 0
 US-09-774-490-1 (1-2709) x B35041 (1-5037)
 QY 64 GGGTCGACACCCCTCGACGTGGCA 41
 Db 2269 GlySerThrProLeuAspValAla 2276
 RESULT 57
 T30192
 probable peptide synthetase - Aureobasidium pullulans
 C;Species: Aureobasidium pullulans
 C;Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 17-Nov-2000
 C;Accession: T30192
 R;Peery, R.B.; Thornnewell, S.J.; Tobin, M.B.; Skatrud, P.I.
 submitted to the EMBL Data Library, January 1997
 A;Description: Discovery of an MDR-like gene adjacent to a peptide synthetase in Aureoba
 A;Reference number: Z20767
 A;Accession: T30192
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: DNA
 A;Residues: 1-10797 <PEE>
 A;Cross-references: EMBL:U85909; NID:g4099310; PID:g4099313; PIDN:AAD00581.1
 C;Genetics:
 A;Introns: 2078/3; 4142/3; 6075/3; 7963/3; 7985/2
 C;Superfamily: acyl carrier protein homology
 C;Keywords: carrier protein
 P;1618-1688/Domain: acyl carrier protein homology <ACP1>

F:3682-3752/Domain: acyl carrier protein homology <ACP2>
 F:5615-5685/Domain: acyl carrier protein homology <ACP3>
 F:7503-7573/Domain: acyl carrier protein homology <ACP4>
 F:9683-9752/Domain: acyl carrier protein homology <ACP5>

Alignment Scores:
 Pred. No.: 25.7 Length: 10797
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservativeness: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x T30192 (1-10797)

QY 645 ATCTGAGGACATATTTTAAAGC 668

Db 1781 lleleuArgThrilePheleuser 1788

RESULT 58

S09700

phycobiliprotein 18.3 beta chain - Synechococcus sp. (PCC 6301) (fragment)

C:Species: Synechococcus sp.

A:Variety: PCC 6301

C:Date: 21-Nov-1993 #sequence_revision 01-Mar-1996 #text_change 09-May-1997

R:Accession: S09700

R:Lundell, D.J.; Glazer, A.N.

J. Biol. Chem. 258, 894-901, 1983

A:Title: Molecular architecture of a light-harvesting antenna. Structure of the 18 S cod

A:Reference number: S09700; PMID:83108897; PMID:6401720

A:Accession: S09700

A:Molecule type: protein

A:Residues: 1-16 <LUN>

A:Experimental source: PCC 6301

Alignment Scores:
 Pred. No.: 630 Length: 16
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservativeness: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.80% Indels: 0
 DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S09700 (1-16)

QY 738 AGGATGCTGTGACGAGCTTA 718

Db 2 ArgAspAlaValSerSerLeu 8

RESULT 59

B40256

interleukin-7 receptor, soluble form - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 23-May-1997

R:Accession: B40256

R:Pfeiman, C.M.; Gimpel, S.D.; Park, L.S.; Hazada, H.; Taniguchi, T.; Ziegler, S.F.

Mol. Cell. Biol. 11, 3052-3059, 1991

A:Title: Organization of the murine and human interleukin-7 receptor genes: two mRNAs ge

A:Reference number: A40256; PMID:91246172; PMID:2038316

A:Accession: B40256

A:Molecule type: DNA

A:Residues: 1-22 <PLE>

A:Note: the authors translated the codon GAA for residue 20 as Gln

C:Superfamily: interleukin-7 receptor; fibronectin type III repeat homology

C:Keywords: cytokine receptor

Alignment Scores:
 Pred. No.: 608 Length: 22
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservativeness: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.80% Indels: 0
 DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x B40256 (1-22)

QY 918 TCCTCAGGATGTCACCTCTCT 898

Db 4 SerSerGlyLeuSerLeuSer 10

RESULT 60

B82805

hypothetical protein XF0448 [imported] - Xylella fastidiosa (strain 9a5c)

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000

C:Accession: B82805

R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; PMID:20365717; PMID:10910347

A:Note: for a complete list of authors see reference number A59328 below

A:Accession: B82805

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-35 <SIM>

A:Cross-references: GB:AE003895; GB:AE003849; NID:9105283; PIDN:AP83258.1; GSPDB:GN001

A:Experimental source: Strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Al

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H.

as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Prohm

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.B.; Laig

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.

, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawaak

A:Authors: da Silva, A.C.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir

M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z

A:Reference number: A59328

A:Contents: annotation

C:Genetics:

A:Gene: XF0448

Alignment Scores:
 Pred. No.: 578 Length: 35
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservativeness: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.80% Indels: 0
 DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x B82805 (1-35)

QY 402 CTACTCCGTTCTCATCCAA 382

Db 15 LeuLeuArgSerSerSerLys 21

RESULT 61

S03674

photosystem I protein psaj - Euglena gracilis chloroplast

C:Species: chloroplast Euglena gracilis

C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 26-Aug-1999

C:Accession: S03674; S26069

R:Manzara, T.; Hallick, R.B.

Nucleic Acids Res. 16, 9866, 1988

A:Title: Nucleotide sequence of the Euglena gracilis chloroplast genes for serine and pr

A:Reference number: S03674; PMID:89041581; PMID:3141903

A:Accession: S03674

A:Molecule type: DNA

A:Residues: 1-37 <MAN>

A:Cross-references: EMBL:M18672; NID:G336891; PIDN:AAA84231.1; PID:G896264

C:Genetics:

A:Gene: psaj

A:Genome: chloroplast

C:Superfamily: photosystem I protein psaj

C:Keywords: chloroplast; photosystem I; transmembrane protein

Alignment Scores:
Pred. No.: 575 Length: 37
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S03674 (1-37)

QY 725 CTGACAGCATCCCTTTTAATA 745

DB 21 LeuThrAlaSerLeuLeuIle 27

RESULT 62

S05490

alpha-amylase (EC 3.2.1.1) 2.46 precursor - wheat (fragment)

C:Species: Triticum aestivum (common wheat)

C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 06-Dec-1996

C:Accession: S05490

R:Huttlly, A.K.; Martienssen, R.A.; Baulcombe, D.C.

Mol. Gen. Genet. 214, 232-240, 1988

A:Title: Sequence heterogeneity and differential expression of the alpha--Amy-2 gene fam

A:Reference number: S05486; UID:89181522; PMID:2467183

A:Accession: S05490

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-51 <STI>

A:Cross-references: EMBL:X13579

C:Genetics:

A:Gene: amy2

A:Map position: 7A

A:Introns: 14/1

C:Function:

A:Description: catalyzes the hydrolysis of internal 1,4-alpha-D-glucosidic bonds

A:Pathway: glycohen/starach degradation

C:Superfamily: wheat alpha-amylase; alpha-amylase core homology

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Alignment Scores:

Pred. No.: 555 Length: 51
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S05490 (1-51)

QY 179 CTTTCTGCTTTAGTCCTTCCT 159

DB 8 LeuCysCysPheSerLeuPro 14

RESULT 63

D42194

pepH protein - Synecococcus sp. (PCC 7002) (fragment)

C:Species: Synecococcus sp.

C:Date: 04-Mar-1993 #sequence_revision 02-Aug-1994 #text_change 26-Aug-1999

C:Accession: D42194

R:Schluchter, W.M.; Bryant, D.A.

Biochemistry 31, 3092-3102, 1992

A:Title: Molecular characterization of ferredoxin-NADP+ oxidoreductase in cyanobacteria:

A:Reference number: A42194; UID:92207922; PMID:1554697

A:Accession: D42194

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-63 <SCH>

A:Cross-references: GB:M86234; GB:J05366; NID:g154535; PIDN:AAA27325.1; PID:g154538

A:Experimental source: PCC 7002

A:Note: sequence extracted from NCBI backbone (NCBI:91780, NCBIP:91792)

C:Superfamily: ompR protein; response regulator homology

F:21-63/Domain: response regulator homology (fragment) <RRH>

Alignment Scores:
Pred. No.: 542 Length: 63
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x D42194 (1-63)

QY 110 GCTGACATGGGAGACAGGC 90

DB 48 AlaAspAsnGlyArgThrGly 54

RESULT 64

S33581

hypothetical protein 3 - Azorhizobium caulinodans

C:Species: Azorhizobium caulinodans

C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 11-May-2000

C:Accession: S33581

R:Stigter, J.; Schneider, M.; de Bruijn, P.J.

Mol. Plant Microbe Interact. 6, 238-252, 1993

A:Title: Azorhizobium caulinodans nitrogen fixation (nif/fix) gene regulation: mutagenes

A:Reference number: S33579; UID:93229807; PMID:8471796

A:Accession: S33581

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-68 <STI>

A:Cross-references: EMBL:X69959; NID:g311387; PIDN:CAA49583.1; PID:g311390

Alignment Scores:

Pred. No.: 537 Length: 68
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x S33581 (1-68)

QY 1719 CTGGGGTTGCCAGCTCCCTT 1739

DB 24 LeuGlyLeuProSerSerLeu 30

RESULT 65

A83261

hypothetical protein PA3085 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: A83261

R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Ba

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path

A:Reference number: A82950; UID:20437337; PMID:10984043

A:Accession: A83261

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-87 <STO>

A:Cross-references: GB:AE004732; GB:AE004091; NID:g9949186; PIDN:AAG06473.1; GSPDB:GN001

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA3085

Alignment Scores:

Pred. No.: 523 Length: 87
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0

DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x A83261 (1-87)
QY 284 CATTGTTCTCCCATTCATTCGAT 264
DB 67 HiscysSerSerHisSerAsp 73
RESULT 66
T25449
hypothetical protein B0412.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Jun-2000
C:Accession: T25449
R:Bentley, D.
A:Submitted to the EMBL Data Library, December 1996
A:Description: The sequence of C. elegans cosmid B0412.
A:Reference number: Z20037
A:Accession: T25449
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-88 <SEN>
A:Cross-references: EMBL:U80953; PIDN:AAB52557.1; GSPDB:GN00021; CESP:B0412.4
A:Experimental source: strain Bristol N2; clone B0412
C:Genetics:
A:Gene: CESP:B0412.4
A:Map position: 3
A:Introns: 4/3; 28/2; 61/3
C:Superfamily: Escherichia coli ribosomal protein S14
Alignment Scores:
Pred. No.: 522 Length: 88
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x T25449 (1-88)
QY 2406 CTTGCGCGACGTGTTTCGG 2386
DB 45 LeuCysArgArgcysPheArg 51
RESULT 67
T30626
hypothetical protein 24L - Molluscum contagiosum virus 1
N:Alternate names: MC024L
C:Species: Molluscum contagiosum virus 1
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 11-May-2000
C:Accession: T30626
R:Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.
A:Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re
A:Reference number: Z20876; MUID:96325459; PMID:8670425
A:Accession: T30626
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-88 <SEN>
A:Cross-references: EMBL:U60315; PIDN:AAC55152.1
C:Genetics:
A:Note: MC024L
Alignment Scores:
Pred. No.: 522 Length: 88
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x T30626 (1-88)

QY 40 ATGCCACCTCGAGGGTGTCTCG 60
DB 38 MetProArgArgGlyValser 44
RESULT 68
I68530
hemoglobin alpha chain (clone alphaG-28) - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 19-May-2000
C:Accession: I68530
R:Lam, V.M.; Gu, Y.L.; Au, D.M.; Wong, W.M.; Ma, C.W.; Cheng, L.Y.
A:Title: Two new rat alpha-globin sequences as identified by the conserved region PCR.
A:Reference number: T54239; MUID:94042225; PMID:8226096
A:Accession: I68530
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-89 <RES>
A:Cross-references: GB:S66658; NID:9439804
C:Genetics:
A:Introns: 32/2
C:Superfamily: globin; globin homology
C:Keywords: blood; oxygen carrier
F:59/Binding site: oxygen (His) (distal axial ligand) #status predicted
Alignment Scores:
Pred. No.: 522 Length: 89
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x I68530 (1-89)
QY 1464 ACTATTGGCGCATTTTCATA 1444
DB 29 ThrileGlyArgLeuPheille 35
RESULT 69
AF2068
hypothetical protein asl2100 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AF2068
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ara
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AF2068
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-91 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA873799.1; PID:G17131191; GSPDB:GN00179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: asl2100
Alignment Scores:
Pred. No.: 520 Length: 91
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 2 Gaps: 0
US-09-774-490-1 (1-2709) x AF2068 (1-91)
QY 1969 ATTCTCTACACCATAGATGAT 1949
DB 63 IleLeuTyThrIleAspAsp 69

RESULT 70
AE2336
potassium-dependent ATPase chain G [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AE2336
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AE2336
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-98 <KUR>
A:Cross-references: GB:BA000019; PIDN:BAB75943.1; PID:g17133379; GSPDB:GN001179
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: asl4244

Alignment Scores:
Pred. No.: 516 Length: 98
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
Gaps: 0
DB:

US-09-774-490-1 (1-2709) x AE2336 (1-98)

QY 2127 ATTCAGCAATACCTCTGCC 2147
Db 17 IIEGINAlailethrSeraLa 23

RESULT 71
G84410
hypothetical protein Vng2613h [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon, K.H.; Alam, M.; Freitas, T.
Proc Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li, A.; Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: G84410
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-99 <STO>
A:Cross-references: GB:AE004437; NID:g10581994; PIDN:AAG20651.1; GSPDB:GN001138
C:Genetics:
A:Gene: VNG2613H

Alignment Scores:
Pred. No.: 516 Length: 99
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
Gaps: 0
DB:

US-09-774-490-1 (1-2709) x G84410 (1-99)

QY 233 TGGGAGTATTACTTACAGCA 253
Db 89 TrpGlyValLeuLeuThraLa 95

RESULT 72
A69514
hypothetical protein AF2113 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C:Accession: A69514
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.P.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson, J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.; Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Uutterback, T.; Cotton, M.D.; Spriggs, T.; Artiaich, P.; Kaine, B.P.; Sykes, S.; Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeon Pyrococcus furiosus
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: A69514
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-99 <KLE>
A:Cross-references: GB:AE000958; GB:AB000782; NID:g2689281; PIDN:AAB89154.1; PID:g2648483

Alignment Scores:
Pred. No.: 516 Length: 99
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
Gaps: 0
DB:

US-09-774-490-1 (1-2709) x A69514 (1-99)

QY 2210 GAGCATTGGGAACCTCTT 2230
Db 6 GluHisLeuGluLeuLeu 12

RESULT 73
AD2566
hypothetical protein alr8527 [imported] - Nostoc sp. (strain PCC 7120) plasmid pCC7120c6
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Dec-2002
C:Accession: AD2566
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. strain PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AD2566
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-100 <KUR>
A:Cross-references: GB:AF003604; PIDN:BAB77446.1; PID:g17134890; GSPDB:GN001183
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: alr8527
A:Genome: plasmid

Alignment Scores:
Pred. No.: 515 Length: 100
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0
DB:

US-09-774-490-1 (1-2709) x AD2566 (1-100)

QY 2517 GCTCAGACATCTCTGGTGCC 2497
Db 46 AlaGlnThrLeuLeuGlyAla 52

RESULT 74
D70408
conserved hypothetical protein aq_1254 - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 20-Sep-1999

Search completed: August 3, 2003, 10:34:52
Job time : 186.5 secs

C:Accession: D70408
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320
A:Accession: D70408
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-102 <AQF>
A:Cross-references: GB:AE000731; NID:g2983691; PIDN:AAC07266.1; PID:g2983702; GB:AE00065
A:Experimental source: strain VF5
C:Genetics:
A:Gene: aq_1254
C:Superfamily: yajC protein

Alignment Scores:
Pred. No.: 514 Length: 102
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x D70408 (1-102)

QY 1404 AAGTTTAAATCATCATCAGGA 1384
|||||
Db 51 LysValIleThrSerGly 57

RESULT 75

IPXL2

insulin II precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C>Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 16-Jul-1999
C:Accession: B33847; S13537
R:Shuldiner, A.R.; Phillips, S.; Roberts Jr., C.T.; LeRoith, D.; Roth, J.
J. Biol. Chem. 264, 9428-9432, 1989
A:Title: Xenopus laevis contains two nonallelic preproinsulin genes. cDNA cloning and ev
A:Reference number: A33847; MUID:89255444; PMID:2722842
A:Accession: B33847
A:Molecule type: mRNA
A:Residues: 1-106 <SHU1>
A:Cross-references: GB:M24442; GB:J04796; NID:g214534; PIDN:AAA49887.1; PID:g214535
R:Shuldiner, A.R.; Bennett, C.; Robinson, E.A.; Roth, J.
Endocrinology 125, 469-477, 1989
A:Title: Isolation and characterization of two different insulins from an amphibian, Xen
A:Reference number: S07199; MUID:89289601; PMID:2661211
A:Accession: S13537

A:Molecule type: protein
A:Residues: 24-53;86-106 <SHU2>
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-53/Domain: insulin chain B #status experimental <BCH>
F:24-53;86-106/Product: insulin #status experimental <MAT>
F:56-83/Domain: connecting peptide #status predicted <CEP>
F:86-106/Domain: insulin chain A #status experimental <ACH>
F:30-92,42-105,91-96/Disulfide bonds: #status predicted

Alignment Scores:

Pred. No.: 512 Length: 106
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x IPXL2 (1-106)

QY 1977 GTGCTACTATCTCTACACCA 1957

|||||
Db 13 ValLeuLeuPheSerThrPro 19

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 15:53:26 ; Search time 597 Seconds
(without alignments)
9361.255 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709

Sequence: 1 aatctttattttatcgatg.....aggcttttttctctaataacc 2709

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1439767 seqs, 1031500376 residues

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq.*
- 3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq.*
- 7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq.*
- 10: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq.*
- 11: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq.*
- 12: /cgn2_6/prodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq.*
- 14: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq.*
- 15: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq.*
- 16: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq.*
- 17: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2709	100.0	2709	9	US-09-774-490-1
2	2508.4	92.6	2530	14	Sequence 1, Appli
3	2508.4	92.6	2530	14	Sequence 283, App
4	2508.4	92.6	2530	14	Sequence 9, Appli
5	718.6	26.5	3871	11	US-10-097-340-283
6	718.6	26.5	3871	11	US-10-262-538-9
7	718.6	26.5	3871	13	US-09-946-374-309
8	718.6	26.5	3871	13	US-10-052-586-347
9	718.6	26.5	3871	14	US-10-174-590-347
10	718.6	26.5	3871	14	US-10-176-758-347
11	718.6	26.5	3871	14	US-10-175-737-347
12	718.6	26.5	3871	14	US-10-173-706-347
13	718.6	26.5	3871	14	US-10-175-738-347
14	718.6	26.5	3871	14	US-10-175-752-347
15	718.6	26.5	3871	14	US-10-176-482-347
16	718.6	26.5	3871	14	US-10-176-757-347
17	718.6	26.5	3871	14	US-10-176-913-347
18	718.6	26.5	3871	14	US-10-180-552-347
19	718.6	26.5	3871	14	US-10-180-557-347

ALIGNMENTS

RESULT 1

US-09-774-490-1
; Sequence 1, Application US/09774490
; Patent No. US20010034332A1
; GENERAL INFORMATION:
; APPLICANT: Jin, Shengfang
; TITLE OF INVENTION: RESISTANCE SEQUENCES AND USES THEREOF
; FILE REFERENCE: 07334-138001
; CURRENT APPLICATION NUMBER: US/09/774,490
; CURRENT FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: US 60/179,191
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2709
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-774-490-1

Query Match	100.0%	Score 2709;	DB 9;	Length 2709;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2709;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AATCTTTTATTTTATCGATGTTAAACAAGCTTAGTAATCGATGCCACGTCGAGGGGTGCG	60	
Db	1	AATCTTTTATTTTATCGATGTTAAACAAGCTTAGTAATCGATGCCACGTCGAGGGGTGCG	60	
Qy	61	ACCCACGGTCCGGGAGTAGGTTGAGCTCGCGCTGTTCTCCCATTCCTCAGCCAGTCTATTT	120	
Db	61	ACCCACGGTCCGGGAGTAGGTTGAGCTCGCGCTGTTCTCCCATTCCTCAGCCAGTCTATTT	120	
Qy	121	CCAGATTGTTTGAACCTTCTCTGGCCGCAATACAGGAAGGAAGACTAAACAGCAAGG	180	
Db	121	CCAGATTGTTTGAACCTTCTCTGGCCGCAATACAGGAAGGAAGACTAAACAGCAAGG	180	
Qy	191	GACCTACAGCGCTGCGAGCATGGGCTGGTGAAGTAGGTTGCTGCTCTTTCTGGGGAGT	240	
Db	191	GACCTACAGCGCTGCGAGCATGGGCTGGTGAAGTAGGTTGCTGCTCTTTCTGGGGAGT	240	

Db 2401 GCAAAGGCCAGACATACCCAGGGAACAGTAAACAATGGAAGCACTTACAAGAAATAA 2460
Qy 2461 GAAAGGTAGAAACAGGAGGACCCAGAAATTTGAGAGGGGACCCAGGAGTGTCTGAGCTGC 2520
Db 2461 GAAAGGTAGAAACAGGAGGACCCAGAAATTTGAGAGGGGACCCAGGAGTGTCTGAGCTGC 2520
Qy 2521 ATTACTCTAGAAACCTCAACAAGTAGAGAACTTGCCTAGACAATACTGGAAAAACAAA 2580
Db 2521 ATTACTCTAGAAACCTCAACAAGTAGAGAACTTGCCTAGACAATACTGGAAAAACAAA 2580
Qy 2581 TCAATATACATGAACCTTTTTCATGGCAATATGAGGATTTTACAATGGTGGAAATTC 2640
Db 2581 TCAATATACATGAACCTTTTTCATGGCAATATGAGGATTTTACAATGGTGGAAATTC 2640
Qy 2641 AGCTGAGTTCACCAATTATAAATTAATCCATGAGTAACTTCTTAATAGGCTTTTTTT 2700
Db 2641 AGCTGAGTTCACCAATTATAAATTAATCCATGAGTAACTTCTTAATAGGCTTTTTTT 2700
Qy 2701 CCTAATACC 2709
Db 2701 CCTAATACC 2709

RESULT 2
US-10-097-340-283
; Sequence 283, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNVARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 283
; LENGTH: 2530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-097-340-283

Query Match 92.6%; Score 2508.4; DB 14; Length 2530;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 193 CTGACGATGGCTGGTTAACTAGGATTTGTCTCTTTCTGGGAGTATTACTTTACAGC 252
Db 9 CTGACGATGGCTGGTTAACTAGGATTTGTCTCTTTCTGGGAGTATTACTTTACAGC 68
Qy 253 AAGAGCAAACTATCAGAAATGGGAAGAAACAATGTGCAAGGCTGAAATTTATCTCAAGA 312
Db 69 AAGAGCAAACTATCAGAAATGGGAAGAAACAATGTGCAAGGCTGAAATTTATCTCAAGA 128
Qy 313 AATGTTGGAATCCCAACATGTGATCATTCAATGGCTTGGCCACACAGCTCCAGTTATCA 372
Db 129 AATGTTGGAATCCCAACATGTGATCATTCAATGGCTTGGCCACACAGCTCCAGTTATCA 188
Qy 373 TACCTTCCCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATAT 432
Db 189 TACCTTCCCTTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATAT 248
Qy 433 TTCAATCGACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGCGCCAGTATCTTACAC 492
Db 249 TTCAATCGACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGCGCCAGTATCTTACAC 308
Qy 493 CAGAAGAGATGAATGCAAGTGGCTGGAAAAAGACATCTCTGAAAGATGTGCTAATTTTCA 552
Db 309 CAGAAGAGATGAATGCAAGTGGCTGGAAAAAGACATCTCTGAAAGATGTGCTAATTTTCA 368
Qy 553 CAAGGTACTTAAAGGCATATAATCAGACTCATTGTGACCCCTGTGAAACGGGGCTTTTCA 612
Db 369 CAAGGTACTTAAAGGCATATAATCAGACTCATTGTGACCCCTGTGAAACGGGGCTTTTCA 428
Qy 613 TCCAAATTTGACCTACATTGAAATTTGGAACATCATCTGAGGACAAATATTTTAAAGCTGA 672
Db 429 TCCAAATTTGACCTACATTGAAATTTGGAACATCATCTGAGGACAAATATTTTAAAGCTGA 488
Qy 673 GAACTCACAATTTGAAACCGCCGCTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGC 732
Db 489 GAACTCACAATTTGAAACCGCCGCTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGC 548
Qy 733 ATCCCTTTTAAATAGATGAGAAATATATCTCTGGAATCTGAGTCTGATTTTATGGGCGGAGA 792
Db 549 ATCCCTTTTAAATAGATGAGAAATATATCTCTGGAATCTGAGTCTGATTTTATGGGCGGAGA 608
Qy 793 CTTTGCTATCTTCCGAACTCTTGGGCAACCCACCACCAATCAGGACAGAGCATGATTC 852
Db 609 CTTTGCTATCTTCCGAACTCTTGGGCAACCCACCACCAATCAGGACAGAGCATGATTC 668
Qy 853 CAGTGGCTCAATGATCAAAAGTTTCAATAGTGGCCACCTCATCTCAGAGAGTGAACAATCC 912
Db 659 CAGTGGCTCAATGATCAAAAGTTTCAATAGTGGCCACCTCATCTCAGAGAGTGAACAATCC 728
Qy 913 TGAAGATGACAAAGTATATCTTTCTCCGTGAAATCAATAGATGAGAACTCTCTGG 972
Db 729 TGAAGATGACAAAGTATATCTTTCTCCGTGAAATCAATAGATGAGAACTCTCTGG 788
Qy 973 AAAAGCTACTCACGCTAGAATAGGTTCAGATATGCAAGAAATGACCTTTGGAGGSCACAGAAG 1032
Db 789 AAAAGCTACTCACGCTAGAATAGGTTCAGATATGCAAGAAATGACCTTTGGAGGSCACAGAAG 848
Qy 1033 TCTGGTGAATTAATGGACAACTTCTCAAAAGCTGCTGATTTGCTCAGTGGCCAGGTCC 1092
Db 849 TCTGGTGAATTAATGGACAACTTCTCAAAAGCTGCTGATTTGCTCAGTGGCCAGGTCC 908
Qy 1093 AAATGGCATTGACACTCATTTTGTATGAACTGACAGATGATTTCTTAATGAATTTTAAAGA 1152
Db 909 AAATGGCATTGACACTCATTTTGTATGAACTGACAGATGATTTCTTAATGAATTTTAAAGA 968
Qy 1153 TCCTAAAAATCCAGTTGTATATGAGTGTTTTACGACTTCCAGTAACATTTTCAAGGGATC 1212
Db 969 TCCTAAAAATCCAGTTGTATATGAGTGTTTTACGACTTCCAGTAACATTTTCAAGGGATC 1028
Qy 1213 AGCCGCTGTATGATATAGCATGATGTGAGAAAGGTGTCTTGGTGCATATATGCCCA 1272

Db 1029 AGCCGTGTGTATGTATAGCATGATGTGAGAGGGTGTCTCTTGGTCCATATGCCCA 1088
Qy 1273 CAGGGATGACCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTCTATCCAGGCC 1332
Db 1089 CAGGGATGACCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTCTATCCAGGCC 1148
Qy 1333 AGGAATCTGTCCAGCAAAACATTTGGTGGTTTGACTCTCAAAAGGACCTTCTGTATGA 1392
Db 1149 AGGAATCTGTCCAGCAAAACATTTGGTGGTTTGACTCTCAAAAGGACCTTCTGTATGA 1208
Qy 1393 TGTATTAACCTTTGCAAGAGTCATCCAGCCATGTACAATCAGTGTTCCTATGAACAA 1452
Db 1209 TGTATTAACCTTTGCAAGAGTCATCCAGCCATGTACAATCAGTGTTCCTATGAACAA 1268
Qy 1453 TCGCCCAATAGTATCAAAACGATGTAAATTTATCAAAATTTGCTGTAGACCG 1512
Db 1269 TCGCCCAATAGTATCAAAACGATGTAAATTTATCAAAATTTGCTGTAGACCG 1328
Qy 1513 AGTGGATGCAAGATGGAACAGTATGTATTTATCGGAACAGATGTTGGACCGT 1572
Db 1329 AGTGGATGCAAGATGGAACAGTATGTATTTATCGGAACAGATGTTGGACCGT 1388
Qy 1573 TCTTAAAGTATCTCAATCTTAAGAGACTTGGTATGATTTAGAGAGGTTCTGTGGA 1632
Db 1389 TCTTAAAGTATCTCAATCTTAAGAGACTTGGTATGATTTAGAGAGGTTCTGTGGA 1448
Qy 1633 AGAATGACAGTCTTTCGGGAACCGACTCTATTTTCAGCAATGGAGCTTTCACATAAGCA 1692
Db 1449 AGAATGACAGTCTTTCGGGAACCGACTCTATTTTCAGCAATGGAGCTTTCACATAAGCA 1508
Qy 1693 GCAACAACTATATTTGGTTCACCGCTGGGGTGCACAGCTCCCTTTACACCGGTGGA 1752
Db 1509 GCAACAACTATATTTGGTTCACCGCTGGGGTGCACAGCTCCCTTTACACCGGTGGA 1568
Qy 1753 TATTTACGGGAAGGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTCTGGGA 1812
Db 1569 TATTTACGGGAAGGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTCTGGGA 1628
Qy 1813 TGGTCTGCTATGCTCTGCTATTTCCCACTCAAGAGACGCAAGACGACAAGATAT 1872
Db 1629 TGGTCTGCTATGCTCTGCTATTTCCCACTCAAGAGACGCAAGACGACAAGATAT 1688
Qy 1873 AAGAAATGGAGACCACTGACTCAGTGTTCAGACTTTACACATGATTAATCACCATGGCCA 1932
Db 1689 AAGAAATGGAGACCACTGACTCAGTGTTCAGACTTTACACATGATTAATCACCATGGCCA 1748
Qy 1933 CAGCCCTGAAGAGAGATCATCTATGGTGTAGAGATAGTAGACATTTTGGAAATGCG 1992
Db 1749 CAGCCCTGAAGAGAGATCATCTATGGTGTAGAGATAGTAGACATTTTGGAAATGCG 1808
Qy 1993 TCCGAAGTCGAGAGAGCGCTGTCTATTGGCAATTCAGAGCGCAAAATGAAGACGAAA 2052
Db 1809 TCCGAAGTCGAGAGAGCGCTGTCTATTGGCAATTCAGAGCGCAAAATGAAGACGAAA 1868
Qy 2053 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGCGCTTCTGTACGTAG 2112
Db 1869 AGAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGCGCTTCTGTACGTAG 1928
Qy 2113 TCTACAAAGAGAGATCAGGCAATTAATCTGTCCGATGGGTGGAAACATGGGTTCATACA 2172
Db 1929 TCTACAAAGAGAGATCAGGCAATTAATCTGTCCGATGGGTGGAAACATGGGTTCATACA 1988
Qy 2173 AACTCTTCTTAAGGTAAACCTCGAAGTCAATTCAGACAGAGATTTGGGAAGACTTCTTCA 2232
Db 1989 AACTCTTCTTAAGGTAAACCTCGAAGTCAATTCAGACAGAGATTTGGGAAGACTTCTTCA 2048
Qy 2233 TAAAGATGATGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2292
Db 2049 TAAAGATGATGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACCTAG 2108
Qy 2293 CCAGAGGTCTGTGACAGAGACTTCATGAGCTTCATCAACACCCCAATCTCAACAGAT 2352
Db 2109 CCAGAGGTCTGTGACAGAGACTTCATGAGCTTCATCAACACCCCAATCTCAACAGAT 2168

Qy 2353 GGATGAGTTCGTGAAACAAAGTTTGGAAAGGACCGCAAAACAAACCTCGCAAGGCCAGG 2412
Db 2169 GGATGAGTTCGTGAAACAAAGTTTGGAAAGGACCGCAAAACAAACCTCGCAAGGCCAGG 2228
Qy 2413 ACATACCCCGAGGAACAGTAACAAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAAA 2472
Db 2229 ACATACCCCGAGGAACAGTAACAAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAAA 2288
Qy 2473 CAGGAGGACCCACGAATTTGAGAGGCGCACCCAGGAGTGTCTGAGCTGCATTACCTCTAGA 2532
Db 2289 CAGGAGGACCCACGAATTTGAGAGGCGCACCCAGGAGTGTCTGAGCTGCATTACCTCTAGA 2348
Qy 2533 AACCTCAAAACAGTAGAAACTTGCCTAGACAATAACTGGAATAACAAATGCAATATACAT 2592
Db 2349 AACCTCAAAACAGTAGAAACTTGCCTAGACAATAACTGGAATAACAAATGCAATATACAT 2408
Qy 2593 GAACCTTTTTCATGGCATTATGTGATGTTTACAATGGTGGAAATTCAGCTGAGTTCCA 2652
Db 2409 GAACCTTTTTCATGGCATTATGTGATGTTTACAATGGTGGAAATTCAGCTGAGTTCCA 2468
Qy 2653 CCAATTATAAATTAATCCATGAGTAACTTTCCTAATAGGCTTTTTC 2702
Db 2469 CCAATTATAAATTAATCCATGAGTAACTTTCCTAATAGGCTTTTTC 2518

RESULT 3

US-10-262-538-9

; Sequence 9, Application US/10262538

; Publication No. US2003011324A1

; GENERAL INFORMATION:

; APPLICANT: Alicalto et al

; TITLE OF INVENTION: NEUROFILIN/VEGF-C/VEGFR-3 MATERIALS AND METHODS

; FILE REFERENCE: 28967/37564

; CURRENT APPLICATION NUMBER: US/10/262,538

; CURRENT FILING DATE: 2002-09-30

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 9

; LENGTH: 2530

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (16)..(2331)

US-10-262-538-9

Query Match 92.6%; Score 2508.4; DB 14; Length 2530;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 2509; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 193 CTGCAGCATGGCTGGTTAACTAGGATTTGTCTGCTTTTCTGGGAGATTACTTTACAGC 252
Db 9 CTGCAGCATGGCTGGTTAACTAGGATTTGTCTGCTTTTCTGGGAGATTACTTTACAGC 68
Qy 253 AAGAGCAACTATCAGAAATGGGAAGAACAAATGTGCCAAGGCTGAAATTTCTTACAAAGA 312
Db 69 AAGAGCAACTATCAGAAATGGGAAGAACAAATGTGCCAAGGCTGAAATTTCTTACAAAGA 128
Qy 313 AATGTTGGAATCCAAATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCA 372
Db 129 AATGTTGGAATCCAAATGTGATCACTTTCAATGGCTTGGCCAAACAGCTCCAGTTATCA 188
Qy 373 TACCTTCCTTTTGGATGAGGAACCGAGTAGCTGTATGTTGGAGCAAGGATCACATATT 432
Db 189 TACCTTCCTTTTGGATGAGGAACCGAGTAGCTGTATGTTGGAGCAAGGATCACATATT 248
Qy 433 TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGTGGCCAGATCTTTACAC 492
Db 249 TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGTGGCCAGATCTTTACAC 308
Qy 493 CAGAAGAGATCAATGCAAGTGGCTGGAAAGACATCTCTGAAAGATGTGCTAATTTAT 552

Db 309 CAGAGAGATGATGCAAGTGGCTGGAAGACATCTGGAAGAAATGTCTAAATTTCA 368
Qy 553 CAAGGTACTTAAGGCATATAATCAGACTCACTTGTACGCTGTGGAACGGGGCTTTTCA 612
Db 369 CAAGGTACTTAAGGCATATAATCAGACTCACTTGTACGCTGTGGAACGGGGCTTTTCA 428
Qy 613 TCCAAATTCGACCTACATTTGAAATTTGGACATCATCTGAGGACAATATTTTAAAGCTGGA 672
Db 429 TCCAAATTCGACCTACATTTGAAATTTGGACATCATCTGAGGACAATATTTTAAAGCTGGA 488
Qy 673 GAACTCACATTTTGAAGAACGGCGTGGGAAGAGTCCATATGACCTTAAGCTGCTGACGC 732
Db 489 GAACTCACATTTTGAAGAACGGCGTGGGAAGAGTCCATATGACCTTAAGCTGCTGACGC 548
Qy 733 ATCCCTTTTAAATAGATGGAAGAAATTAATCTCTGGAACCTGACGCTGATTTTATGGGGCGAGA 792
Db 549 ATCCCTTTTAAATAGATGGAAGAAATTAATCTCTGGAACCTGACGCTGATTTTATGGGGCGAGA 608
Qy 793 CTTTGTCTATCTCCGAACTCTTGGGCAACCAACCCCAATCAGGACAGACGACATGATTC 852
Db 609 CTTTGTCTATCTCCGAACTCTTGGGCAACCAACCCCAATCAGGACAGACGACATGATTC 668
Qy 853 CAGGTGGCTCAATGATCCAAAGTTTCAATAGTGCACCTCATCTCAGAGAGTGACAATCC 912
Db 669 CAGGTGGCTCAATGATCCAAAGTTTCAATAGTGCACCTCATCTCAGAGAGTGACAATCC 728
Qy 913 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGGAAGAACACTCTGG 972
Db 729 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAATGCAATAGATGGAAGAACACTCTGG 788
Qy 973 AAAAGTACTCAGCTAGATAGTGCAGATATGCAAGATGACTTTGGAGGGGACAGAG 1032
Db 789 AAAAGTACTCAGCTAGATAGTGCAGATATGCAAGATGACTTTGGAGGGGACAGAG 848
Qy 1033 TCTGTGTAATAATGGAACAACATCTCTCAAGCTCGTCTGATTTGCTCAGTGCAGAGTCC 1092
Db 849 TCTGTGTAATAATGGAACAACATCTCTCAAGCTCGTCTGATTTGCTCAGTGCAGAGTCC 908
Qy 1093 AAATGGCATGTGACATCAATTTTGTGATGAACTGCAGAGTGATTTCTTAATGAACTTTAAAGA 1152
Db 909 AAATGGCATGTGACATCAATTTTGTGATGAACTGCAGAGTGATTTCTTAATGAACTTTAAAGA 968
Qy 1153 TCTTAATAATCCAGTTGATATGAGTGTGTTAGACTTCCAGTCAATTTTCAAGGGATC 1212
Db 969 TCTTAATAATCCAGTTGATATGAGTGTGTTAGACTTCCAGTCAATTTTCAAGGGATC 1028
Qy 1213 AGCCGTGTGTATGTATAGCATCAGTGTGATGAGAGGGGTGTTCTTGGTCCATATGCCCA 1272
Db 1029 AGCCGTGTGTATGTATAGCATCAGTGTGATGAGAGGGGTGTTCTTGGTCCATATGCCCA 1088
Qy 1273 CAGGATGGAACCAACTATCAATGGGTGCTTTATCAAGGAAGTCCCTATCCAGGCTC 1332
Db 1089 CAGGATGGAACCAACTATCAATGGGTGCTTTATCAAGGAAGTCCCTATCCAGGCTC 1148
Qy 1333 AGGAATTTGCTCCAGCAAAACATTTGGTGGTGTGATCTCAAAAGGACCTTCTGATGA 1392
Db 1149 AGGAATTTGCTCCAGCAAAACATTTGGTGGTGTGATCTCAAAAGGACCTTCTGATGA 1208
Qy 1393 TGTTATACTTTGCAAGAGTCATCCAGCCATGTACAAATCAGTGTCTTCTATGAACAA 1452
Db 1209 TGTTATACTTTGCAAGAGTCATCCAGCCATGTACAAATCAGTGTCTTCTATGAACAA 1268
Qy 1453 TCGCCCAATGATGATCAAAACGGATGTAAATTTATCAATTTACAAATTTGCTAGACCG 1512
Db 1269 TCGCCCAATGATGATCAAAACGGATGTAAATTTATCAATTTACAAATTTGCTAGACCG 1328
Qy 1513 AGTGAATGCAAGATGGAAGTATGATGTTATGTTTATCGGAAACAGATTTGGGACCGT 1572
Db 1329 AGTGAATGCAAGATGGAAGTATGATGTTATGTTTATCGGAAACAGATTTGGGACCGT 1388
Qy 1573 TCTTAAAGTATTTCAATTTCTTAAGGACCTTGGTATGATTTAGAGAGGTTCTCTGGA 1632
Db 1389 TCTTAAAGTATTTCAATTTCTTAAGGACCTTGGTATGATTTAGAGAGGTTCTCTGGA 1448

Qy 1633 AGAAATGACAGTTTTTTCGGGAACCGACTGCTATTTTTCAGCAATGGAGCTTTTCCACTAAGCA 1692
Db 1449 AGAAATGACAGTTTTTTCGGGAACCGACTGCTATTTTTCAGCAATGGAGCTTTTCCACTAAGCA 1508
Qy 1693 GCAACAACTATATATTTGGTTCAACGGCTGGGGTTCGCCAGCTCCCTTTTACACGGGTGGA 1752
Db 1509 GCAACAACTATATATTTGGTTCAACGGCTGGGGTTCGCCAGCTCCCTTTTACACGGGTGGA 1568
Qy 1753 TATTTACGGGAAGCGTGTGCTGAGTGTTCCTCCCGAGACCTTACTGTGCTTGGGA 1812
Db 1569 TATTTACGGGAAGCGTGTGCTGAGTGTTCCTCCCGAGACCTTACTGTGCTTGGGA 1628
Qy 1813 TGGTCTGCTATGTTCTGCTATTTTCCCACTGCAAGAGACGCAACAAGACGACAAGATAT 1872
Db 1629 TGGTCTGCTATGTTCTGCTATTTTCCCACTGCAAGAGACGCAACAAGACGACAAGATAT 1688
Qy 1873 AAGAAATGGAGACCCACTGACTCACTGTTTCCAGACTTACCAATGATATCACCANTGGCCA 1932
Db 1689 AAGAAATGGAGACCCACTGACTCACTGTTTCCAGACTTACCAATGATATCACCANTGGCCA 1748
Qy 1933 CAGCCCTGAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGACAATTTTTCGAAATGCAG 1992
Db 1749 CAGCCCTGAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGACAATTTTTCGAAATGCAG 1808
Qy 1993 TCCGAAGTCGAGAGAGCGCTGGTCTATTGGCAATTTCCAGAGGGGAAATGAAGAGCGAAA 2052
Db 1809 TCCGAAGTCGAGAGAGCGCTGGTCTATTGGCAATTTCCAGAGGGGAAATGAAGAGCGAAA 1868
Qy 2053 AGAAGAGATCAGAGTGGATGATCATCATCAGGACAGATCAAGGCCCTTCTGCTACGTAG 2112
Db 1869 AGAAGAGATCAGAGTGGATGATCATCATCAGGACAGATCAAGGCCCTTCTGCTACGTAG 1928
Qy 2113 TCTCAACAGAGAGGATTCAGGCAATTTACTCTGCCATCGGTGGGAACATGGTTCATACA 2172
Db 1929 TCTCAACAGAGAGGATTCAGGCAATTTACTCTGCCATCGGTGGGAACATGGTTCATACA 1988
Qy 2173 AACTCTTCTTAAGGTAACCTTGGAAAGTCAITTGACAAGAGCAATTTGGGAAGCACTTCTTCA 2232
Db 1989 AACTCTTCTTAAGGTAACCTTGGAAAGTCAITTGACAAGAGCAATTTGGGAAGCACTTCTTCA 2048
Qy 2233 TAAAGATGATGATGAGATGAGCTCTAAGACCAAGAAATGTCCTAATAGATGACACCTAG 2292
Db 2049 TAAAGATGATGATGAGATGAGCTCTAAGACCAAGAAATGTCCTAATAGATGACACCTAG 2108
Qy 2293 CCAGAGGTCTGTACAGAGACTTTCATGAGTCTCAACCAACCCCAATCTCAACAGCAT 2352
Db 2109 CCAGAGGTCTGTACAGAGACTTTCATGAGTCTCAACCAACCCCAATCTCAACAGCAT 2168
Qy 2353 GGATGAGTCTGTGAACAAAGTTTGGAAAGGGACCGAAAAACAACGTCGGCAAGGCCAGG 2412
Db 2169 GGATGAGTCTGTGAACAAAGTTTGGAAAGGGACCGAAAAACAACGTCGGCAAGGCCAGG 2228
Qy 2413 ACATACCCCGGGAACAGTAACTAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAAA 2472
Db 2229 ACATACCCCGGGAACAGTAACTAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAAA 2288
Qy 2473 CAGGAGGACCCAGCAATTTTGGAGGGCACCCAGAGGTGCTGAGCTGCAATTTACCTCTAGA 2532
Db 2289 CAGGAGGACCCAGCAATTTTGGAGGGCACCCAGAGGTGCTGAGCTGCAATTTACCTCTAGA 2348
Qy 2533 AACCTCAAAACAGTAGAACTTGGCTAGACAATTAACCTGAAAAAACAATGCAATATACAT 2592
Db 2349 AACCTCAAAACAGTAGAACTTGGCTAGACAATTAACCTGAAAAAACAATGCAATATACAT 2408
Qy 2593 GAACTTTTTCATGGCAATTTATGAGTGTGGAATTTTACAAATGGTGGGAATTCAGCTGAGTTCCA 2652
Db 2409 GAACTTTTTCATGGCAATTTATGAGTGTGGAATTTTACAAATGGTGGGAATTCAGCTGAGTTCCA 2468
Qy 2653 CCAATTAATAATTAATCCAGTAACTTTTCCCTAAATAGGCTTTTTCCTC 2702
Db 2469 CCAATTAATAATTAATCCAGTAACTTTTCCCTAAATAGGCTTTTTCCTC 2518

RESULT 4
US-09-946-374-309
Sequence 309, Application US/09946374
Publication No. US20030073129A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnovers, Luc
APPLICANT: Eaton, Dan L.
APPLICANT: Ferrara, Napoleone
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, Christopher J.
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K.
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2830PIC1
CURRENT APPLICATION NUMBER: US/09/946,374
CURRENT FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: 60/098716
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098723
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098749
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098750
PRIOR FILING DATE: 1998-09-01
PRIOR APPLICATION NUMBER: 60/098803
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/098821
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/098843
PRIOR FILING DATE: 1998-09-02
PRIOR APPLICATION NUMBER: 60/099536
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099596
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099598
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099602
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099642
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099741
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099754
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099763
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099792
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099808
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099815
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099816
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100385
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100388
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100390
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100584
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100627
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100661
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100662
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100664
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100683
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100684
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100710
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100711
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100848
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100849
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100919
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/100930
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101014
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/101068
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/101071
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/101279
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: 60/101471
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101472
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101474
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101475
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101476
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101477
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101479
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101738
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101741
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101743
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101915
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101916
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/102207
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102240
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102307
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102330
PRIOR FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 60/102331
PRIOR FILING DATE: 1998-09-29

;; PRIOR APPLICATION NUMBER: 60/102484
;; PRIOR FILING DATE: 1998-09-30
;; PRIOR APPLICATION NUMBER: 60/102487
;; PRIOR FILING DATE: 1998-09-30
;; PRIOR APPLICATION NUMBER: 60/102570
;; PRIOR FILING DATE: 1998-09-30
;; PRIOR APPLICATION NUMBER: 60/102571
;; PRIOR FILING DATE: 1998-09-30
;; PRIOR APPLICATION NUMBER: 60/102684
;; PRIOR FILING DATE: 1998-10-01
;; PRIOR APPLICATION NUMBER: 60/102687
;; PRIOR FILING DATE: 1998-10-01
;; PRIOR APPLICATION NUMBER: 60/102965
;; PRIOR FILING DATE: 1998-10-02
;; PRIOR APPLICATION NUMBER: 60/103258
;; PRIOR FILING DATE: 1998-10-06
;; PRIOR APPLICATION NUMBER: 60/103314
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103315
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103328
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103395
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103396
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103401
;; PRIOR FILING DATE: 1998-10-07
;; PRIOR APPLICATION NUMBER: 60/103449
;; PRIOR FILING DATE: 1998-10-06
;; PRIOR APPLICATION NUMBER: 60/103633
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/103678
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/103679
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/103711
;; PRIOR FILING DATE: 1998-10-08
;; PRIOR APPLICATION NUMBER: 60/104257
;; PRIOR FILING DATE: 1998-10-14
;; PRIOR APPLICATION NUMBER: 60/104987
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105000
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105002
;; PRIOR FILING DATE: 1998-10-20
;; PRIOR APPLICATION NUMBER: 60/105104
;; PRIOR FILING DATE: 1998-10-21
;; PRIOR APPLICATION NUMBER: 60/105169
;; PRIOR FILING DATE: 1998-10-22
;; PRIOR APPLICATION NUMBER: 60/105266
;; PRIOR FILING DATE: 1998-10-22
;; PRIOR APPLICATION NUMBER: 60/105693
;; PRIOR FILING DATE: 1998-10-26
;; PRIOR APPLICATION NUMBER: 60/105694
;; PRIOR FILING DATE: 1998-10-26
;; PRIOR APPLICATION NUMBER: 60/105807

Query Match 26.5%; Score 718.6; DB 11; Length 3871;

Best Local Similarity 61.9%; Pred. No. 7.9e-205; Mismatches 739; Indels 21; Gaps 5;

Matches 1233; Conservative 0;

Qy 269 AATGGGAAGAACATGTGCGAAGGCTGAAATATCTCTACAAAGAAATGTGGAAATCCCAAC 328
Db 215 ACTTTGAAGCAAAATATTCGAAGACTCAAGCTTAACCTACAAAGACTTGTCTCTTCAAT 274

Qy 329 AATGTGATCACTTCAATGGCTTGGCCCAACAGCTCCAGCTTATCATPACCTTCTCTTTGGAT 388
Db 275 AGCTGTATTCCTTTTGGGTTTCATCAGAAGGACTGGATTTCAAACTCTTCTCTTAGAT 334

Qy 389 GAGGAACGAGTAGGCTGTATGTTGGCAAGAGTACATATTTTCATTTCGACCTGGTT 448
Db 335 GAGGAAGAGGAGGCTGCTCTTGGGAGCCAAAGACCATCTTTCTACTCAGTCTGGTT 394

Qy 449 AA---TATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACACAGAGAGATGAA 505
Db 395 GACTTAAACAAAAATTTTAAAGAAGATTTATTTGGCCCTGTGCAAGGAACGGGTGAAATTA 454

Qy 506 TGCAGTGGGCTGGAAAAGACATCCTGAAAGATGTCTTAATTTTCATCAAGTACTTAAAG 565
Db 455 TGTAAATTTAGCTGGGNAAGATGCCAATACAGAAATGTGCAAAATTTTCATCAGAGTACTTCAG 514

Qy 566 GCATATAATCAGACTCACTGTAGCCCTGTGGAACGGGGCTTTTCATCCCAATTTGCAACC 625
Db 515 CCCTATAACAAACTCACAATATATGTGTGGAAGTGGAGCATTTTCATCCCAATATGTGGG 574

Qy 626 TACATTTGAAATTTGGACATCATCTTGAGGACAATATTTTAAAGCTGGAGAACTCACATTTT 685
Db 575 TATATTGATCTTGGAGTCTACAAGGAGGATATATATTTCAAACTAGACACACATAATTTG 634

Qy 686 GAAAACGGCCGTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGATCCCTTTTAATA 745
Db 635 GAGTCTGGCAGACTGAAATGTCTTTCGATCCTCAGCAGCCCTTTTGTCTCAGTAAATGACA 694

Qy 746 GATGAGAAATATATCTCTGGAACCTGCAGCTGATTTTATGGGCGAGACTTTTGTCTATCTTC 805
Db 695 GATGAGTACCTCTACTCTGGAACAGCTTCTGATTTCTTGGCAAGAGTACTGCAATTCAT 754

Qy 806 CGAACTCTTGGGC-----ACCACCAACCAATCAGACAGAGCAGCATGATTTCCAGG 856
Db 755 CGATCCCTTGGGCCCTACTCATGACCACCACTACATCAGAACTGACATTTTCAGAGCACTAC 814

Qy 857 TGGCTCAATGATCCAAAGTTCAATTAGTGCCCACTCATCTCAGAGAGTGACATCTCTGAA 916
Db 815 TGGCTCAATGAGCAAAATTTTATTTGGAACCTTCTTCATACCAGACACCTACAATCCAGAT 874

Qy 917 GATGACAAAGTATATCTTTTCTCGTGAAATGCAATAGATGGAGAACACTCTGGAATA 976
Db 875 GATGATAAAATATATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 934

Qy 977 GCTACTCAGCTAGATAGTGCAGATATGCAAGAAATGATTTTGGAGGCGACAGAGTCTG 1036
Db 935 ACCATCTCTCTCGAGTTGGAAGAGTTTGTAAAGATGATGTAGGAGGACAAACGCGACCTG 994

Qy 1037 GTGAATAATGGACAACTTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCCTCAAT 1096
Db 995 ATAAACCAAGTGGACGACTTTTCTTAAAGCCAGACCTGATTTTCTCAATTTCTCTGGAAGT 1054

Qy 1097 GGCATTGACACTCATTTTGTGAACTGCAGAGATGATTTCTTAAAGAACTTTTAAAGATCCT 1156
Db 1055 GGGCAGATATCTTCTTGTGAGCTTCAAGATATTTTATTTACTCCCAAGAGATGAA 1114

Qy 1157 AAAAATCCAGTTGATATGAGTGTTTACGACTTTCCAGTAACTTTTCAAGGGATCAGCC 1216
Db 1115 AGAAATCCTGTAGTATATGAGTCTTTTACTACAACCACTGCTCAATTTTCAAGGGCTCTGCT 1174

Qy 1217 GTGTGATGATAGCATGAGTGTGAGAAGGGTGTCTTGTGCTCATATGCCACAGG 1276
Db 1175 GTTTGTGTATAGCATGGCTGACATCAGAGCAGTGTTTTAAATGGTCTCATATGCTCATAG 1234

Qy 1277 GATGACCCCACTATCAATGGTGCCTTTATCAAGGAAGAGTCCCTATCCACGGCCAGGA 1336
Db 1235 GAAAGTGCAGACCACTGTTGGGTGCAGTATGATGGGAGAAATTCCTTATCCAGGGCTGGT 1294

Qy 1337 ACTTGTCCCAAGCAAAACATTTG---GTGGTTTGTGACTCTTACAAAGGACCTTTCTCTGATGAT 1393
Db 1295 ACATGTCCAAGCAAAACCTATGACCCCACTGATTAAGTCCACCCGAGATTTTCCAGATGAT 1354

Qy 1394 GTTATAACCTTTGCAAGAGTCTCCAGCCATGTAGCAATCCAGTGTGTTTCTCTATGAAACAT 1453
Db 1355 GTCATCATGTTTCAATAAGCGGCACTCTGTGATGTATAAGTCCGTATATACCCAGTTGCGAGA 1414

Qy 1454 CGCCCAATAGTATCAAAACGGATGTAATTTATCAATTTTACACAAATTTGCTGAGACCGA 1513
Db 1415 GGACCAACGTTCAAGAGAAATCAATGTGGATTTACAGACTGACACAGATAGTGGTGGATCAT 1474

;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/086023
;; PRIOR FILING DATE: 1998-05-18
;; PRIOR APPLICATION NUMBER: 60/086392
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/086486
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/087098
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087208
;; PRIOR FILING DATE: 1998-05-28
;; PRIOR APPLICATION NUMBER: 60/087609
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087759
;; PRIOR FILING DATE: 1998-06-02
;; PRIOR APPLICATION NUMBER: 60/087827
;; PRIOR FILING DATE: 1998-06-03
;; PRIOR APPLICATION NUMBER: 60/088025
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088028
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088029
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088033
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088167
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088202
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088212
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088217
;; PRIOR FILING DATE: 1998-06-05
;; PRIOR APPLICATION NUMBER: 60/088326
;; PRIOR FILING DATE: 1998-06-04
;; PRIOR APPLICATION NUMBER: 60/088655

;; PRIOR FILING DATE: 1998-06-09
;; PRIOR APPLICATION NUMBER: 60/088722
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088738
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088740
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088811
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088824
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088825
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088826
;; PRIOR FILING DATE: 1998-06-10
;; PRIOR APPLICATION NUMBER: 60/088861
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088863
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/088876
;; PRIOR FILING DATE: 1998-06-11
;; PRIOR APPLICATION NUMBER: 60/089090
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089105
;; PRIOR FILING DATE: 1998-06-12
;; PRIOR APPLICATION NUMBER: 60/089512
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089514
;; PRIOR FILING DATE: 1998-06-16
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089598
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089653
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089908

Query Match 26.5%; Score 718.6; DB 13; Length 3871;
Best Local Similarity 61.9%; Pred. No. 7.9e-205;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;

QY 269 AATGGGAAGAACAAATGTGCGAAGGCTGAAATATCTCAAGAAATGTGGAATCAAC 328
DB 215 ACTTTGAAGCAAAATATTTCCAAGACTCAAGCTAACCTACAAAGACTTGTCTTTCAAT 274
QY 329 AATGTGATCATTCTCAATGGCTTGGCCAAAGCTCCAGTTATCATACCTTCTCTTTGGAT 388
DB 275 AGCTGTATTCCCTTTTGGGTTTCATCAGAAGACTGGGATTTTCAAACCTCTTCTTAGAT 334
QY 389 GAGGAACGGAGTAGGCTGTATGTTGGAGCAAGGATCATATATTTTCATTCGACCTGGTT 448
DB 335 GAGGAAGAGCGAGGCTGCTCTTGGAGCCAAAGACCACATCTTCTACTCAGTCUGTT 394
QY 449 AA---TATCAAGGATTTTCAAAAAGATTGTGGCCAGTATCTTACCAAGAGAGATGAA 505
DB 395 GACTTAAACAAAAATTTTAAGAAGATTTATTGGCCCTGCTGCAAGGAACGGTGAATTA 454
QY 506 TCGAAGTGGCTGGAAAGACATCTGAAAGATGCTGCTTAATTTTCATCAAGTACTTAG 565
DB 455 TGTAATTTAGCTGGGAAGAGTCCCAATACAGATGTGCAAAATTTTCATCAGAGTACTTCAG 514
QY 566 GCATATAATCAGACTCACTTGTACGCTGTGGAACGGGGCTTTTTCATCCAAATTTGCACC 625
DB 515 CCCTATACAAACTCATATATGTGTGGAAGTGGAGCATTTTCATCCCAATATGTGG 574
QY 626 TACATTGAAATTTGGACATCATCTGAGGACAAATATTTTTTAAGCTGGAGAACTCAATTT 685
DB 575 TATATTGATCTTGGAGTCTACAAGAGGATATTATATTTCAAACCTAGACACACATAATTG 634
QY 686 GAAAACGGCCGTGGGAAGAGTCCCATATGACCTTAAGCTGCTGACAGCATCCCTTTTAATA 745
DB 635 GAGTCTGCCAGACTGAATGTCTTTCATCTCCAGCAGCCTTTTGGTTCAGTAATGACA 694

Qy	389	GAGAA	CGGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATATTTTCATTTCGACCTGGTT	448
Db	335	GAGAA	AGAGCGAGCGCTGCTCTTGGGAGCCAAAGACACATCTTTCTACTCAGTCTGGTT	394
Qy	449	AA---	TATCAGGATTTTCAAAAGATTGTGTGGCCAGTATCTTACACCGAGAGATGAA	505
Db	395	GACTTAA	CAAAAATTTTAAAGAAATTTATTGGCCCTGCTGCAAGGAAACGGGTGGAATTA	454
Qy	506	TGCAAGTGGG	TGAAAAAGACATCCTGAAAGAAATGTGCTAAATTTTCATCAAGGTACTTTAAG	565
Db	455	TGTAATATTAGCT	TGGAAAGATGCCAATACAGAAATGTGCAAAATTTTCATCAGAGTACTTCAG	514
Qy	566	GCATATAATCAGACT	CACTTGTACGCTGTGGAACGGGGCTTTTCATCCAAATTTGCAACC	625
Db	515	CCCTATAACAAAAT	CTCATATATGTGTGGAACCTGAGCATTTTCATCCAATATGTGGG	574
Qy	626	TACATTGAAAT	TGACATCATCTCGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTT	685
Db	575	TATATTGATCT	TGGAGTCTCAAGAGGAGATATATTTCAAACTAGACACACATAAATTTG	634
Qy	686	GA AAAACGGCCGT	TGGAAAGAGTCCATATGACCCCTAAGCTGTGCAGAGCATCCCTTTTAATA	745
Db	635	GAGTCTGCAGACT	GAAATGTCTTTTTCGATCCTCAGCAGCCTTTTGCTTCAGTAATGACA	694
Qy	746	GATGGAAAT	TATACTCTGGAACCTGCAGCTGAATTTTATGGGGCGAGACTTTGCTATCTTC	805
Db	695	GATGAGTACCTCT	ACTCTGGAACAGCTTCGTATTTCTTGGCAAGATACTGCAATTCAC	754
Qy	806	CGAACTCTTGGGC	-----ACCACCAACCAATCAGGACAGGACGATGATTCAGG	856
Db	755	CGATCCCTTGGG	CCCTACTCATGACCCACCACTACATCAGAACTGACATTTTCAGAGCACTAC	814
Qy	857	TGGCTCAATGAT	CCAAAAGTTTCATTAGTGGCCCACTCATCTCAGAGAGTGACAATCTGAA	916
Db	815	TGGCTCAATGGAGCA	AAATTTATTGGAACTTTCTTCATACAGACACTACAATCCAGAT	874
Qy	917	GATGACAAAAGT	ATACTTTTCTCCGTGAAAATGCAATAGATGAGAACTCTGGAAAA	976
Db	875	GATGATAAAAAT	ATATTTCTTCTTTCGTGAAATCATCTCAAGAAAGCAGTACTCCGATAAA	934
Qy	977	GCTACTCACGCT	AGAAATAGTTCAGATATGCAGAAATGACTTTGGAGGGCAAGAAHCTG	1036
Db	935	ACCATCCTTTCT	CGAGTTGGAAGAGTTGTGAAGAATGATGTAGGAGGACAAACGACGCTG	994
Qy	1037	GTGAATAAAT	TGGACAACTTCTCAAAAGCTCGTCTGATTTGCTCAGTGGCCAGGTTCCAAT	1096
Db	995	ATAAACAGTGG	ACGACTTTTCTTAAAGCCAGACTGATTTGCTCAATTTCTTGGAAHGTAT	1054
Qy	1097	GGCAATGACACT	CAATTTTGATGAACCTGCAGGATGATTTCTTAATGAACCTTTAAAGATCCT	1156
Db	1055	GGGCAGATACT	TACTTTGATGAGCTTCAAGATATTTATTACTCCCAACAGAGATGAA	1114
Qy	1157	AAAATCCAGTT	TGTATATGAGHGTTTACGACTTCCAGTAACTATTTCAAGGATCAGCC	1216
Db	1115	AGAAATCCTGT	AGTATATGAGTCTTTTACTACAAACCACTCCATCTTTCAAAGGCTCTGCT	1174
Qy	1217	GTGTGTATGAT	AGCATGATGTCAGAAAGGAGTTCCTTGTGTCCATATGCCACAGG	1276
Db	1175	GTTTGTGTAT	AGCATGGCTGACATCAGAGCAGTTTTAAATGGTCCATATGCTCATAG	1234
Qy	1277	GATGGACCCAA	CTATCAATGGTGGCTTTATCAAGGAAGAGTCCCTTATCCACGGCCAGGA	1336
Db	1235	GAAAGTCAGAC	CACTCGTTGGTGCAGTATGATGGGAGAAATTCCTTATCCACGGCCTGGT	1294
Qy	1337	ACTTGTCCAG	CAAAAATTTG---GTGGTTTGGACTCTCAAAAGGACCTTCTCTGATGAT	1393
Db	1295	ACATGTCCAA	AGCAAAACCTATGACCCCACTGATTAAAGTCCACCCGAGATTTTCCAGATGAT	1354
Qy	1394	GTTATAACCT	TTTGCAGAAGTTCATCCAGCCATGTACAAATCCAGTGTTCCTATGAAACAAT	1453
Db	1355	GTCAATCAGT	TTTCATAAAGCGGCATCTGTGATGTATAAGTCCGATATACCAGTTGACGA	1414
Qy	1454	CGCCCAAT	TAGTATCAAAAACGGATGTAATTTATCAATTTTACACAAAATGTGCTGAGACCGA	1513

Db	1415	GGACCAACGTTTCAAGAGAACTCAATGTGGATTACAGACTGCACACAGATAGTGGTGGATCAT	1474
Qy	1514	GTGATGCGAGAGATGACAGTAGTATGTTATCTTTTTCGGAACAGATGTTGGACCGGTT	1573
Db	1475	GTCAATGCGAAGATGCCAGTAGTACGATGTAAATGTTTCTTGGAAACAGACATTTGGAACGTGC	1534
Qy	1574	CTTAAAGTAGTTTCAATTTCTTAAGGAGACTGGTATGATTAGAAAGGTTCTGCTGGAA	1633
Db	1535	CTCAAGTTGTCAAGCTTTCAAGGAAAAGTGG---AATATGAGAGGAGTAGTCTGGAG	1591
Qy	1634	GAATGACAGTTTTTCGGGAACCGACTGCTATTTCACGAATGAGCTTTCCACTAAGCAG	1693
Db	1592	GAGTTGCAGATATTCAAGCACTCATCAATCATCTTGAACATGAAATGTTCTCTGAAGCAG	1651
Qy	1694	CAACAACTATATTGGTTCAACGGCTGGGGTTGCCAGCTCCCTTTTACACCGGTGTGAT	1753
Db	1652	CAACAAATTGACATTTGGTTCCCGAGATGGAATTAGTTCAGCTCTCTTTGCAAGATCGAC	1711
Qy	1754	ATTTACGGGAAAAGCGTGTGCTGAGTGTGCTCGCCGAGACCTTACTGTCTGGGAT	1813
Db	1712	ACTTATGGGAAAAGCTTGCAGACACTGTTGTCTTGGCAGAGACCCCTACTGTGCTGGAT	1771
Qy	1814	GGTTCTGCATGTTCTCGCTATTTTCCCACTGCAAAAGAGACGCAAGACATATA	1873
Db	1772	GGAAATGCATGCTCTCGATATGCTCTACTTCTTAAAGGAGAGCTAGACGCCAAGATGA	1831
Qy	1874	AGAAATGGAGACCGCACTGACTCATGTTTCAGACTTACACCATGATTAATCACCATGGCCAC	1933
Db	1832	AAATATGGCGACCCCAATCACCCAGTGTGGGACATCGAAGACAGCATTAGTCTATG---AA	1888
Qy	1934	AGCCCTCAAGAGAGATCATCTATGGTGTAGAAATAGTAGACATTTTGGAAATGCAGT	1993
Db	1889	ACTGCTGATGAAGAAGTGATTTTTTGGCAATTGAAATTTAACTCAACCTTTCTGGAATGTATA	1948
Qy	1994	CCGAAGTCGAGAGAGCGCTGCTATATTGGCAATTTCCAGAGCGCAATCAAGAGCGCAAA	2053
Db	1949	CCTAAATCCCAACAAGCAACTATTAAATGGTATATCCAGAGGTGAGGGATGAGCATCGA	2008
Qy	2054	GAAGAGATCAGATGGATGATCATATCATCAGAGACAGATCAAGGCCCTTCTGCTACGTAGT	2113
Db	2009	GAGGAGTTGAAGCCGATGAAGAATCATCAAAACGGAATATGCGGTACTGATTCGAAGT	2068
Qy	2114	CTACAACAGAAAGGATTCAGGCAATTACCTCTGCCATCGGTGGAAACATGGTTCATACAA	2173
Db	2069	TTGCAGAAGAAGGATTCCTGGGATGTATTCTGCAAGCCCGAGGAGCACATTTTCATCCAC	2128
Qy	2174	ACTCTTCTTAAGGTAACCCCTGGAAAGTCATTGACACAGAGCATTTTGGAAAGAACTTCTTCAT	2233
Db	2129	ACCATAGTGAAGCTGACTTTGAATGTCAATTGAGATGAACAGATGGAATAATCCAGAGG	2188
Qy	2234	AAAGATGATCATG	2246
Db	2189	GCAGAGCATGAGG	2201

RESULT 7

US-101-176-758-347
; Sequence 347, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRAN

;; TITLE OF INVENTION: ACIDS ENCODING THE SAME
;; FILE REFERENCE: P3430RIC104
;; CURRENT APPLICATION NUMBER: US/10/176,758
;; CURRENT FILING DATE: 2002-06-21
;; Prior Application removed - See File Wrapper or Palm
;; NUMBER OF SEQ ID NOS: 612
;; SEQ ID NO 347
;; LENGTH: 3871
;; TYPE: DNA
;; ORGANISM: Homo Sapien
US-10-176-758-347

Query Match 26.5%; Score 718.6; DB 14; Length 3871;
Best Local Similarity 61.9%; Pred. No. 7.9e-205;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;

QY	269	AATGGGAAGAACATGTGCGAAGGCTGAAATATCTACAAAGAAATGTTGGAATCCAAAC	328
DB	215	ACTTTGAAGCAAAATATCCAGACTCAAGCTAAACCTACAAAGACTTGCTGCTTTCAAT	274
QY	329	AATGTCATCACTTTCAATGCTTGGCCAAAGCTCCAGCTTATCAVACCTTCTTTTGGAT	388
DB	275	AGCTGTATTCCTTTTGGGTTTCATCAGAGACTGGATTTTCAAACTCTTCTTAGAT	334
QY	389	GAGGAACGAGTAGGCTGTATGTGGAGCAAAAGGATCAATATTTTCATTCGACCTGGTT	448
DB	335	GAGGAAGAGGCGAGGCTGCTCTTGGGAGCAAAAGACCATCTTCTACTCAGTCTGGTT	394
QY	449	AA---TATCAGGATTTTCAAAAGATTTGAGGAGTATCTTACAGAGAGATGAA	505
DB	395	GACTTAAACAAAATTTTAAAGATTTTATGCGCTGCTGCAAGGAACGGGTGGAATTA	454
QY	506	TGCAAGTGGCTGGAAAGACATCTCTGAAAGATGTCTAAATTTTCATCAAGTACTTAAG	565
DB	455	TGTAATTAAGTGGGAAAGATGCCAATACAGATGTGCAATTTTATCAGAGTACTTCAG	514
QY	566	GCATATAATCAGACTCACTTGACGCTGTGGAAGGGGGCTTTTCATCCAAATTTGCACC	625
DB	515	CCCTATAACAAAACCTCACATATATGTGTGGAACCTGGAGCATTTTCATCCAAATATGTGG	574
QY	626	TACATGAATTTGGACATCATCTGAGGACATATTTTAACTGGAGACTCAATTTT	685
DB	575	TATATTGATCTTGGAGTCTACAGGAGGATATTTATTTCAAACTAGACACACATAATTTG	634
QY	686	GAAGACGGCTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAATA	745
DB	635	GAGCTGGCAGACTGAAATGCTCTTCGATCTCTCAGAGCTTTTGTCTCAGTAATGACA	694
QY	746	GATGGAAATATATCTCTGGAACCTGAGCTGTATTTATGGGCGGAGACTTTGCTATCTTC	805
DB	695	GATGAGTACCTCTACTCTGGAACAGCTTCTGATTTCTTGGCAAGAGACTGCAATTCAT	754
QY	806	CGAATCTTGGGC-----ACCACCAACCAATCAGACAGAGCAGCATGATTTCCAGG	856
DB	755	CGATCCCTTGGGCCTACTCATGACCACCACTACATCAGAACTGACATTTTCAGAGCACTAC	814
QY	857	TGCTCAATGATCCAAAGTTCAATAGTGCCCACTCATCTCAGAGAGTGACAACTCCTGAA	916
DB	815	TGCTCAATGGAGCAAAATTTATTTGGAATTTCTTTCATACACAGACCTCAATCCAGAT	874
QY	917	GATGACAAAGTATATCTTTTCTCGTGAAATGCAATAGATGGAGAACACTCTCGGAAAA	976
DB	875	GATGATAAAATATATTTCTTCTTCTGTAATCATCTCAAGAGGCGAGTACCTCCGATAAA	934
QY	977	GCTACTCAGCTAGATAGTTCAGATATCAGAAATGATTTTGGAGGCGACAGAGTCTG	1036
DB	935	ACCATCTTTCTCGAGTTTGAAGAGTTTGTAAAGATGATGTAGGAGGACAAAGCGAGCTG	994
QY	1037	GTGAATAATGACAACTTCCTCAAGCTTCCTGATTTGCTAGTTCAGTCCAGGCTCCAAT	1096
DB	995	ATAAACAAGTGGACGACTTTTCTTAAGGCCAGACTGATTTGCTCAATTTCTTGGAGTGT	1054
QY	1097	GCATTTGACACTCATTTTGTATGAACTGCGAGGATGATTTCTTAAATGAACTTTAAAGATCCT	1156

DB	1055	GGGCGAGATACCTTACTTTGATCAGCTTCAAGATATTTATTTACTCCCAAGAGATGAA	1114
QY	1157	AAAATCCAGTTGTATATGGAGTGTTTACGACTTCCAGTAAATTTTCAAGGATCAGCC	1216
DB	1115	AGAAATCCTGTAGTATATGGAGTCTTTTACTACAACAGCTCCATCTTCAAAGGCTCTGT	1174
QY	1217	GTGTGTATGTATAGCATGATGTGAGAGGGTGTTCCTTGGTCCATATATCCCAACAGG	1276
DB	1175	GTTTGTGTATAGCATGTGCTGACATCAGAGCAGTTTTTAAATGGTCCATATGCTCATAAG	1234
QY	1277	GATGACCCAACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATATCCAGGCGAGA	1336
DB	1235	GAAAGTGAGACCACTCGTTGGTGCAGTATGATGGGAGAAATTCCTTATCCACGGCTGT	1294
QY	1337	ACTTGTCCCAAGCAAAACATTTG--GTGGTTTGTACTCTACAAGAGACTTCTCTATGAT	1393
DB	1295	ACATGTCCAAAGCAAAACCTATGACCCACTGATTAACTCCACCGAGATTTTCCAGATGAT	1354
QY	1394	GTATAACTTTGCAAGAAAGTCATCCAGCCATGTACAAATCCAGTGTTCCTTATGAACAAT	1453
DB	1355	GTATCATGTTTATTAAGCGGCACCTCTGTGATGTATAGTCCGTATATCCAGTTCAGGA	1414
QY	1454	CGCCCAATAGTGTATCAAAACCGATGTAATTTATCAATTTTACAAATTTGCTGTAGACGA	1513
DB	1415	GGACCAACGTTCAAGAGAAATCAATGTGATTTACAGACTGACACAGATAGTGTGATCAT	1474
QY	1514	GTGGATCGAGAAGATGGACAGTATGATTTATCGGAACAGATGTTGGGACCGTT	1573
DB	1475	GTCTTGGAGAAGATGGCCAGTACGATGATTAATTTCTTGGAAACAGACATTTGGAATCTGT	1534
QY	1574	CTTAAAGTAGTTTCAATTTCTTAAAGAGACTTGGTATGATTTTAGAAGAGGTTCTCTGGAA	1633
DB	1535	CTCAAGTTGTGAGCATTTCAAGGAAAGTGG--AATATGGAAGAGTAGTGTCTGGAG	1591
QY	1634	GAAATGACAGTTTTTGGGAAACCGACTGTCTATTTTACGAAATTTGAGAGCTTTTCCATTAAGCAG	1693
DB	1592	GAGTTGCAGATATTTCAAGCACTCATCAATCATCTTGAACATGGAATTTCTCTGAAGCAG	1651
QY	1694	CAACAACTATATATTTGTTTCAACGGCTGGGTGCCAGCTCCCTTTTACACCGGTGCTAT	1753
DB	1652	CAACAAATTTGATATTTGTTTCCGAGATGGAATAGTTCAGCTCTCTTGCACAGATGCCAC	1711
QY	1754	ATTTACGGGAAAGCGTGTGCTGAGTGTTCCTCCCGCAGACCCCTTACTGTCTTGGAT	1813
DB	1712	ACTTATGGGAAGCTTGGCAGACTGTGTCTTGCAGAGACCCCTACTGTGCTGGAT	1771
QY	1814	GGTTCGTGATTTCTCGTATTTTCCCACTGCAAGAGAGACGCAAGACGACAGATATA	1873
DB	1772	GGAAATGCTCTCTCGATATGCTCTACTTCTTAAAGGAGAGCTAGACGCAAGATGTA	1831
QY	1874	AGAAATGGAGACCCACTGACTGCTTTCAGACTTACACCATGATTAATCACTTCCAGCCAC	1933
DB	1832	AAATATGGGACCCCAATCACTCCAGTGTGGGACATCGAAGACAGCATTTAGTCATG--AA	1889
QY	1934	AGCCCTCAAGAGAGAAATCATCTATGGTGTAGAGAAATAGTAGCACAATTTTGGAAATGCA	1993
DB	1889	ACTGCTGATGAAGAGTGAATTTTGGCATTTGAATTTTAACTCAACCTTCTTGAATGTATA	1948
QY	1994	CCGAAGTCCGAGAGAGCGCTGTCTTATTTGGCAATTTCCAGAGGCGAAATGAAGAGGGA	2053
DB	1949	CCTAAATCCCAACAAAGCAACTATTTAAATGGTATATCCAGAGGTCAGGAGGATGAGCATCGA	2008
QY	2054	GAAGAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTCTCTAGTGTAGT	2113
DB	2009	GAGGAGTTGAAGCCCGATGAAGAAATCATCAAAACGGAATATGGGCTACTGATTCGAAGT	2069
QY	2114	CTACAACAGAGAGGATTCAGGCAATTTACTCTCTGCCATCGGTTGGAAACATGGGTTTCA	2173
DB	2069	TTGCAGAAAGAGGATTTCTGGAGTGTATTTACTGCAAAAGCCAGGAGACACTTTTATCCAC	2128
QY	2174	ACTCTTCTTAAGGTAAACCTCGGAAGTCAATTGACACAGAGCATTTTGGAAAGACTTCTTCAT	2233

Db 2129 ACCATAGTGAAGCTGACTTTGATGTCAATGTAGATGAACAGATGGAATAATACCCAGAGG 2188
QY 2234 AAGATGATGATG 2246
Db 2189 GCAGAGCATGAGG 2201

RESULT 8

US-10-175-737-347
; Sequence 347, Application US/10175737
; Publication No. US20030013153A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C50
; CURRENT APPLICATION NUMBER: US/10/175,737
; Prior Filing DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 347
; LENGTH: 3871
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-175-737-347

Query Match 26.5%; Score 710.6; DB 14; Length 3871;

Best Local Similarity 61.9%; Pred. No. 7.9e-205;

Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;

QY 269 AATGGAAGAACAAATGTGCCAAGGCTGAAATATCTCAAGAGAAATGTGGAATCCAAAC 328
Db 215 ACTTTGAAGCAAAATATTCGAAGACTCAAGCTAAGCTAAGCTAAGCTAAGCTAAGCTAAGCT 274
QY 329 AATGTGATCATTCAATGGCTTGGCCAAAGCTCCAGCTTATCATATCATATCTTCTTTGGAT 388
Db 275 AGCTGTATTCCTTTTGGGTTTCATCAGAGGACTGGATTTCAAACTCTCTCTTAGAT 334
QY 389 GAGGAACGAGTAGGCTGTATGTTGGAGCAAGAGATCAATATTTTCAATTCGACCTGGTT 448
Db 335 GAGGAAGAGGAGGCTGCTCTTGGGAGCCAAAGACCAATCTTCTACTCAGTCTGGTT 394
QY 449 AA--TATCAAGATTTTCAAGAGATTTGTGGCCAGTATCTTACACAGAGAGATGAA 505
Db 395 GACTTAACAAAATTTTAAGAGATTTATTTGGCTGTCTGCAAGGAAGCGGGTGAATTA 454
QY 506 TGCAAGTGGCTGGGAAAGACATCTGAAAGATTTGTGAGGATTTTCAATCAAGTACTTAAG 565
Db 455 TGTAATAGCTGGGAAAGATCCCAATACAGATTTGCAATTTTCAATCAGTACTTCAG 514
QY 566 GCATATATCAGACTCATTGTGCGCTGTGGAGCGGGGCTTTTTCATCCAAATTTGCACC 625
Db 515 CCTTATAACAAAACCTCACATATATGTGTGGAATTTTCAATCCAAATATGTGGG 574
QY 626 TACATTTGAATTTGGAGAGTCCATATGACCCCTAAGCTGTGACAGCATCCCTTTTAATA 745
Db 575 TATATTTGATCTTGGAGTCTCAAGAGGATATTTATTTCAAACTAGACACATATTTG 634
QY 686 GAAACGCGCTGGGAAAGTCCATATGACCCCTAAGCTGTGACAGCATCCCTTTTAATA 745
Db 635 GAGTCTGGCAGACTGAATGTCTTTCGATCTCTGAGAGCCCTTTTGTCTCAGTATGACA 694
QY 746 GATGAGAAATTAATCTCTGGAACCTGAGCTGATTTTATTTGGGCGAGACTTTGCTATCTTC 805

Db 695 GATGAGTACCTCTACTCTGGAACAGCTTCTGATTTCTTGGCAAGATATGCTCATCTCACT 754
QY 806 CGAACTCTTTGGGC-----ACCACCAACCAATCAGGACAGAGCAGCATGATTTCCAGG 856
Db 755 CGATCCCTTTGGGCTTACTCATGACCAACCATATCATCAGAACTGACATTTTCAGAGCACTAC 814
QY 857 TGGCTCAATGATCCAAAGTTTCAATAGTGCCCACTCATCTCAGAGAGTGAACAATCTCTGAAA 916
Db 815 TGGCTCAATGGAGCAAAATTTTATTTGAACTTTCTTTCATACACAGACACCTTACAACTCCAGAT 874
QY 917 GATGACAAAGTATATCTTTTCTTCCGTGAAATGCAATAGATGGAGAGACACTCTCGAAA 976
Db 875 GATGATAAAATATATTTCTTCTTCTGTAATCATCTCAAGAGAGGAGTACCTCTCGATAAA 934
QY 977 GCTACTCAGCTAGATAGGTAGGTAGATATGCAAGATAGCTTTTGGAGGGGACAGAAAGTCTG 1036
Db 935 ACCATCTTCTCGAGTTTGGAGAGGTTTGTAGAAATGATGTAGGAGGACACAGCGAGCTG 994
QY 1037 GTGAATAAATGGACAAACATTTCTCAAAGCTCGTCTGTGATTTGCTCAGTGCCAGGTCCAAAT 1096
Db 995 ATAAACAAGTGGAGGACTTTTCTTAAGGCCAGACTGATTTGCTCAATTTCTTGGAGTGT 1054
QY 1097 GGCATTGACACTCAATTTTGTGAACTGAGAGTATTCCTTAATGAACTTTTAAAGATCTCT 1156
Db 1055 GGGGAGATATCTTACTTTGATGAGCTTCAAGATATTTTACTCTCCCAAGAGATGAA 1114
QY 1157 AAAAATCCAGTTGTATATGAGTGTGTACGACTTCCAGTAACTTTTCAAGGGATCAGCC 1216
Db 1115 AGAAATCCTGTAGTATATGGAGTCTTTTACTCAACAGCTCCATCTTCAAAGGCTCTGCT 1174
QY 1217 GTGTGTATGTATAGCATGATGTGAGAAAGGGTGTCTTGTGTCTCATATGCCACACAGG 1276
Db 1175 GTTGTGTGTATAGCATGCTGACATCAGAGCAGTTTATTTGCTCATATGCTCATAG 1234
QY 1277 GATGAACCCAACTATCAATGGGTGCTTTATCAAGAAAGAGTCCCTATTCACGGGACAGA 1336
Db 1235 GAAAGTGAGAGACCATCTGTTGGGTGAGTATGATGGAGAAATTTCTTATTCACGGGCTGCT 1294
QY 1337 ACTTGTCCCAAGCAAAACATTTG---GTGGTTTTCAGCTCTCAAAAGGACCTTCTCTGATCAT 1393
Db 1295 ACATGTCCAAAGCAAAACCTATGACCCACTGATTTAGTCCACCGAGATTTTCCAGATGAT 1354
QY 1394 GTTATAACCTTTGCAAGAAAGTCAATCCAGCCATGATCAATCCAGTGTCTTATGAACAAT 1453
Db 1355 GTCATCAGTTTCAATAGCGGCACTCTGTGATGTATAGTCCGTATATCCAGTTGACAGA 1414
QY 1454 CGCCCAATAGTGTATCAAAACGAGTGAATTTATCAATTTTACAAATTTGCTGTAGACCGA 1513
Db 1415 GGAACCAAGTTTCAAGAGAAATCAATGTGATTTACAGACTGACACAGATAGTGTGATCAT 1474
QY 1514 GTGGATGAGAGAGTGGACAGTATGATGTTTATCGGAACAGATGTTGGGACCCCTT 1573
Db 1475 GTCATGAGAAAGATGGCCAGTACGATGATGATTTTCTTGAACACAGATTTGGAACCTGTC 1534
QY 1574 CTTAAAGTGTCTTCAATTTCTTAAGAGACTTTGGTATGATTTTAGAAGAGGTTCTGCTGAA 1633
Db 1535 CTCAAGTTGTGACGATTTCAAGGAAAGTGG---AATATGGAAGAGTGTGCTGAG 1591
QY 1634 GAAATGACAGTTTTCGGGAAACCGACTGCTATTTTACGAATTTGAGAGTTCCTCACTAAGCAG 1693
Db 1592 GAGTTGACAGATTTCAAGCACTCATCAATCATCTTGAACATGGAATTTCTCTGGAAGCAG 1651
QY 1694 CAAACAATATATATTTGTTCAACCGCTGGGTTGCCAGGCTCCCTTACACGGGTGTGAT 1753
Db 1652 CAAACAATTTGATGTTGTTCCGAGATAGTATTTAGTGTGCTCTCTGCAAGATGCGAC 1711
QY 1754 ATTTACGGGAAGCGTGTGCTGAGTGTGCTCTCCCGAGAGCCCTTACTGTGCTGGAT 1813
Db 1712 ACTTATGGGAAGCTTGGCAGACTGTGTCTTCCAGAGACCCCTTACTGTGCTGGAT 1771
QY 1814 GGTCTGATGTTCTCGCTATTTTCCCACTGCAAGAGAGCGCAAGACGACAAAGATATA 1873

1772 GGAATGCAATGCTCTCGATATGCTCTACTTCTTAAAGGAGAGCTAGACGCCAAGATGTA 1831 Db
1874 AGAATGGAGACCCACTGCTCACTGCTCAGACTTACACCATGATAATCACCATGCCAC 1933 Qy
1832 AATATGGGACCCCAATCACCAGTCTGGACATCGAAGACAGCAATTTAGTCATG---AA 1888 Db
1934 AGCCCTGAAGAGAGAAATCATCTATGCTGTAGAGAAATAGTACACATTTTTTGGAAATGCAGT 1993 Qy
1889 ACTGCTGATGAAAGGTGATTTTGGCAATTTGAATTTAACTCAACCTTTCTGGAATGTATA 1948 Db
1994 CCGAAGTCCAGAGAGCGGTGCTATTTGGCAATTTCCAGAGCGGAAATGAAGAGCGAA 2053 Qy
1949 CCTAAATCCCAACAGCACTATTAAATGGTATATCCAGAGGTTCAGGGATGAGCATCGA 2008 Db
2054 GAAGAGATCAGAGTGCATCATATCATCAGACAGATCAAGGCCCTTCTGCTAGCTAGT 2113 Qy
2009 GAGGAGTTGAGCCGATGAAGATCATCAAAACGGAATATGGGCTACTGATTTGAAAT 2068 Db
2114 CTACAAAGAGGATTCAGGCAATTAACCTCTGCCATGCGGTGGAAACATGGTTCATACAA 2173 Qy
2069 TTGCAGAAGAGATTTCTGGATGTATTACTGCAAGCCCGAGGACACACTTTTCATCCAC 2128 Db
2174 ACTTCTTAAAGTAACCTCGAAGTCATGACACAGAGCATTTGGAAGAACTTCTTCAT 2233 Qy
2129 ACCATAGTGAAGCTGACTTTGAATGTCTTGAATGTGATGGAATGAACAGATGGAAATACCCAGAGG 2188 Db
2234 AAGATGATGATG 2246 Qy
2189 GCAGAGCATGAGG 2201 Db

RESULT 9

US-10-173-706-347
; Sequence 347, Application US/10173706
; Publication No. US2003002293A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C7
; CURRENT APPLICATION NUMBER: US/10/173,706
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 347
; LENGTH: 3871
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-173-706-347

Query Match 26.5%; Score 718.6; DB 14; Length 3871;
Best Local Similarity 61.9%; Pred. No. 7.9e-205;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;
Qy 269 AATGGAAGAACAAATGTCAGAGCTGAAATTTATCTCAAGAAATGTTGGAATCCAAAC 328
Db 215 ACTTTGAAGCAAAATATTTCAAGACTCAAGCTAACTCAAGAACTTGTGCTTTCAAT 274
Qy 329 AATGTCATCACTTCAATGGCTTGGCCAAACAGCTCCAGCTTATCATACCTTCTTTGGAT 388
Db 275 AGCTGTATTCCTTTTGGGTTTCATCAGAGGACTGGATTTTCAAACTTCTCTTAGAT 334
Qy 389 GAGGAACGGAGTAGGCTGTATGTTGGAGCAAGGATCATATTTTTCATTCGACCTGGTT 448

335 GAGGAAGAGGCGAGCTGCTCTTGGAGCCAAAGACACATCTTTCTACTCAGTCTGGTT 394 Db
449 AA---TATCAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACACAGAGAGATGAA 505 Qy
395 GACTTAAACAAAATTTTAAAGATTTTATTTGGCCCTGCTGCAAGAGAAACGGGTGGAATTA 454 Db
506 TGCAAGTGGGCTGGAAAGACATCTGAAAGAAATGCTGCTAAATTTTCAATCAAGGTACTTAAG 565 Qy
455 TGTAAATAGCTGGGAAGATGCCAATACAGAAATGTCAAATTTTCAATCAGAGTACTTCAG 514 Db
566 GCATATAATCAGACTCACTTGTACGCTGTGGAAACGGGGCTTTTTCATCCAAATTTGCAACC 625 Qy
515 CCTATACAAAACCTCACATATATGTTGTGGAACTGGAGCATTTTATCCCAATATGTGGG 574 Db
626 TACATTTGAATTTGACATCATCTCTGAGGACAAATATTTTAAAGCTGGGAACTCACAATTTT 685 Qy
575 TATATTGATCTTGGAGTCTCAAGAGAGATATTTATTTCAAACTAGACACACATATTTG 634 Db
686 GAAACCGCCCTGGGAAGAGTCCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAAATA 745 Qy
635 GAGTCTGCGACACTGAAATGTCCTTCGATCCTCAGCAGCCTTTTGTCTCAGTAATGACA 694 Db
746 GATGAGAAATTTATCTCTGGAACCTGCACTGATTTTATGGGGCGAGACTTTTGTCTATCTTC 805 Qy
695 GATGAGTACCTCTACTCTGGAAACAGCTTCTGATTTCTTGGCAAGAGTACTGCAATTCAT 754 Db
806 CGAAGCTCTTGGGCTCTCTGAGGACAACTTCTGAGGACAGGACAGCATGATTTCCAGG 856 Qy
755 CGATCCCTTGGGCTCTCTGAGGACAACTTCTGAGGACAGGACAGCATGATTTCCAGG 814 Db
857 TGGCTCAATGATCCAAAGTTCATTTAGTCCCACTCATCTCAGAGAGTGACAACTCCGAA 916 Qy
815 TGGCTCAATGAGCAAAATTTTGTGAACTTCTTTCATACAGACACCTCAATCCAGAT 874 Db
917 GATGACAAAGTATATTTTCTCCGTGAAATGCAATAGATGGAAGAACACTCTTGGAA 976 Qy
875 GATGATAAATATATTTCTCTTCTTCTCAAGAAAGGCACTCTCCGATGATAA 934 Db
977 GCTACTCAGCTAGATAGTGCAGATGCAAGAGTGTGTAAGAAATGATGAGGAGGACAGCAGCTG 1036 Qy
935 ACCATCTCTTCTCGAGTTGGAAAGTGTGTAAGAAATGATGAGGAGGACAGCAGCTG 994 Db
1037 GTGAATAAATGGACAAACATTCCTCAAGCTCTGCTGATTTGCTCAGTCCAGGTCCAAAT 1096 Qy
995 ATAAACAAGTGGACGACTTTTCTTAAGCCAGACTGATTTGCTCAATTTCTTGGAGTGT 1054 Db
1097 GGCATTCAGCTCAATTTTGTGATGAACCTGAGATGTATTCCTTAATGAACCTTTAAAGATCT 1156 Qy
1055 GGGGCGAGATCTTACTTTGATGAGCTTCAAGATATTTATTTACTCCCCCAAGAGATGAA 1114 Db
1157 AAAATCCAGTTGATATGAGTGTGTTACGACTTCCAGTAACTTTTCAAGGGATCAGCC 1216 Qy
1115 AGAAATCCTGTAGTATATGAGTGTCTTACTTCAACAGCTCCATCTTCAAGGGCTCTGCT 1174 Db
1217 GTGTGATGTATAGCATGATGATGTGAGAGGTTGTTCTTGGTTCATATGCCACAGG 1276 Qy
1175 GTTTGTGTATAGCATGGCTGACATCAGAGCAGTTTATGTTGTTCCATATGCTCATAG 1234 Db
1277 GATGGACCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTATATCCAGGGCCAGGA 1336 Qy
1235 GAAAGTCAGACCACTCGTTGGGTGAGTATGATGGAGAAATTCCTTATTCACGGCTGCT 1294 Db
1337 ACTTGTCCAGCAAAACATTTG---GTGGTTTTCAGCTTCAAAAGGACCTTCTCTGATGAT 1393 Qy
1295 ACATGTCAAGCAAAACCTATGACCCACTGATTAAGTCCACCGAGATTTTCCAGATGAT 1354 Db
1394 GTTATAACCTTTGCAAGAAAGTCAATCCAGGATGTATCAATCCAGTGTTCCTTATGAACAA 1453 Qy
1355 GTCATCAGTTTTCATAGCGGCACTCTGTGATGTATAGTCCGTATATCCAGTTGACAGGA 1414 Db
1454 CGCCCAATAGTGTATCAAAACGGATGTAAATTTATCAATTTTACAAATTTGTGTTAGACCGA 1513 Qy

Db 1415 GGACCAACGTTCAAGAGATCAATGTGGATTACAGACTACACAGATAGTGTGATCAT 1474
QY 1514 GTGGATGCAAGAGATGAGCAGATGATGTTATGTTATCGGAACAGATGTTGGGACGTT 1573
Db 1475 GTCAATGCAAGAGATGCGCAGTACGATGTAATGTTTCTTGGAAACAGACATTTGCAACTGTC 1534
QY 1574 CTTAAAGTAGTTTCAATCTTCAAGAGACTTGGTATGATTTAGAGAGGTTCTGCTGAA 1633
Db 1535 CTCAAAGTGTGACGATTTCAAGGAAAGAGTG--AATATGGAAGAGGTAGTGTCTGGAG 1591
QY 1634 GAAATGACAGTGTCTTCCGGAACCGACTGCTATTTTCCAGCAATGGAGCTTCCACTAAGCAG 1693
Db 1592 GAGTGTGACATATTCAGACACTCATCATCTTGAACATGGAATGTCTCTGAAGCAG 1651
QY 1694 CAACAACTATATATTGGTTCAACGCTGGGTTGCCAGCTCCCTTTTACACCGGTGTGAT 1753
Db 1652 CAACAACTGTACATTTGGTTCGAGATGATTTAGTTTCACTCTCTGACAGATGCGAC 1711
QY 1754 ATTTACGGGAAGCGTGTCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTTGGAT 1813
Db 1712 ACTTATGGGAAGCTTGGCAGACTGTTGTCTTGGCAGAGACCCCTACTGTGCTGGAT 1771
QY 1814 GGTCTGTGATGTTCTCGCTATTTTCCACTGCAAGAGAGACGACAAAGACCAAGATATA 1873
Db 1772 GGAATGCAATGCTCTGATATGCTCTACTCTTAAAGAGAGCTTAGACGCCAAGATGA 1831
QY 1874 AGAATGGAGACCCACTGACTCACTGTTTCCAGCTTACACCATGATTAATCACCATGGCCAC 1933
Db 1832 AATATATGGCACCCTAATCACCCAGTGTGGGACATCGAAGACAGCATTAGTCATG--AA 1888
QY 1934 AGCCCTGAAGAGAGATCATCTATGTTGTAGAGATAGTAGACATTTTGGAAATGCAAT 1993
Db 1889 ACTGTGTGAAGAGGTGATTTTGGCAATGGAATTAATCTCAACCTTTCTGGAATGTATA 1948
QY 1994 CGAAGTCCAGAGAGCGTGTCTATTGGCAATTCAGAGCGGAAATGAAGAGGAAA 2053
Db 1949 CCTAATCCCAAGACAACTAATTAATGTTATTCAGAGGTGAGGATGAGCATCGA 2008
QY 2054 GAAGAGATCAGATGGATGATCATATCATCAGACAGATCAAGGCTTCTGCTAGTAGT 2113
Db 2009 GAGGAGTTGAAGCCGATGAAGAGATCATCAAAACGGAATATGGCTACTGATTCGAAT 2068
QY 2114 CTACAACAGAGATTCAGGAAATTTACCTCTGCCATGCGGTGGAAACATGGGTTTCAACA 2173
Db 2069 TTGCAAGAGAGGATTTCTGGATGATTTACTGCAAGGCCAGGAGACACTTTTCAATCCAC 2128
QY 2174 ACTCTTCTTAAGTAACTGCAAGTCAATGACAGAGCAATTTTGAAGAACTTCTTTCAT 2233
Db 2129 ACATAGTGAAGTGAATTTGAATGTCATTTGAATGAATGAATGAATGAATGAATGAATGA 2188
QY 2234 AAAGATGATGATG 2246
Db 2189 GCAGAGCATGAGG 2201

RESULT 10

US-10-175-738-347
; Sequence 347, Application US/10175738
; Publication No. US2003002294A1

GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3430RIC45
; CURRENT APPLICATION NUMBER: US/10/175,738
; CURRENT FILING DATE: 2002-06-19
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 347
; LENGTH: 3871
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-175-738-347

Query Match 26.5%; Score 718.6; DB 14; Length 3871;

Best Local Similarity 61.9%; Pred. No. 7.9e-205;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;

QY 269 AATGGGAAGAACAAATGTGCCAAGGCTGAAATATATCTTACAAAGAAATGTTCGAATCCCAAC 328
Db 215 ACTTTGAAGCAAAATATTTCCAGACTCAGCTAACCTACAAAGACTTCTCTCTTCAAT 274
QY 329 AATGTGATCACTTTCAATGGCTTGGCCAAAGCTCCAGTTATCATACCTTCTTTTGGAT 388
Db 275 AGCTGTATTCCTTTTGGGTTTCATCAGAGGACTGGATTTTCAAACTCTTCTCTTAGAT 334
QY 389 GAGNAACGAGTACGCTGATGTTGGCAAGGATCACATATTTTCAATTCGACTGCTT 448
Db 335 GAGGAAAGAGGACGAGCTGCTCTTGGAGCCAAAGACCAATCTTTTCTACTCAGTCTGTT 394
QY 449 AA---TATCAAGGATTTTCAAAAGATTGTGGCCAGTATCTTACCCAGAGAGATGAA 505
Db 395 GACTTAAACAAAATTTTGAAGATTTATTTGGCTGCTGCAAGGAACGGTGGAAATTA 454
QY 506 TCGAAGTGGGCTGAAAGAGACATCTGAAAGAAATGTGCTAAATTTTCAATCAAGGTACTTAAG 565
Db 455 TGTAAATTAGCTGGGAAAGATGCCAATACAAATGTGCAAAATTTTCAATCAGAGTACTTCAG 514
QY 566 GCATATATCAGACTCACTGTAGCCCTGTGGAGAGGGGCTTTTTCATCCAAATTTGCACC 625
Db 515 CCCTATAACAAAACCTCATAATATGTGTGGAACTGGAGCATTTTTCATCCAAATATGTTGG 574
QY 626 TACATGGAATTTGACATCATCTTGGAGCAATATTTTAAAGCTGGAGAACTCACATTTT 685
Db 575 TATATTGATCTTGGAGCTTACAGAGAGATTTATTTTCAAACTAGACACACATATTTTG 634
QY 686 GAAACCGCCGCTGGGAAGAGTCCAATATACCTTAAGCTGTGACAGCATCCCTTTTAAATA 745
Db 635 GAGCTGCGAGACTGAAATGTCTTCCATCTCAGCAGCCCTTTTGTCTCAGTAATGACA 694
QY 746 GATGAGAAATTAATCTTGGAACTGCACTGATTTTATGGGCGAGACTTTGCTATCTTC 805
Db 695 GATGAGTACCTCTACTCTGGAACAGCTTCTGATTTCTTGGCAAAGATATCTGCAATTCAT 754
QY 806 CGAACTCTTGGGC-----ACCACCCCAATCAGGACAGAGCAGCATGATTTCAGG 856
Db 755 CGATCCCTTGGGCTACTCATGACCCACCTATCATCAGAACTGACATTTTCAGAGCACTAC 814
QY 857 TGGCTCAATGATCCAAAGTTTCAATAGTCCCACTCATCTCAGAGAGTGAACAATCTCGAA 916
Db 815 TGGCTCAATGGAGCAAAATTTATTTGAACTTTCTTTCATACAGACACCTCAATCCAGAT 874
QY 917 GATGACAAAGTATATCTTTTCTTCCGTGAAATGCAATAGATGAGAGAACACTCTGGAATA 976
Db 875 GATGATAAAATATATTTCTTTCTTCTGTAATCATCTCAAGAAAGGCACTACCTCCGATAA 934
QY 977 GCTACTCAGCTAGATAGGTAGATGCAAGAAATGACTTTTGGAGGCGCAGAGGCTG 1036
Db 935 ACCATCTTTCTCGAGTTGGAAGGTTTGAAGAAATGATGAGGAGGACACAGCAGCTG 994
QY 1037 GTGAATAAATGGCAACAATTTCTCAAGCTCGTCTGATTTTGGCTCAGTCCAGGTCCAAAT 1096
Db 995 ATAAACAGTGGACGACTTTTCTTAAGCCAGACTGATTTGCTCAATTTCTTGGAGTAT 1054
QY 1097 GGCATTGACACTCAATTTTGTATGAATCTCAGAGTATTCCTATGAATCTTTAAAGATCCT 1156

Db 695 GATGAGTACTCTACTCTGGAACAGCTTCTGATTTCTTGGCAAGATACTGCATTCCT 754
Qy 806 CGAAGCTTTGGGC-----ACGACCCCAATCAGACAGAGCAGCATGATTCAGG 856
Db 755 CGATCCCTTTGGGCTACTCATGACCACTACATACAGAACTGACATTTTCAGAGCACTAC 814
Qy 857 TGGCTCAATGATCCAAAGTTTCATAGTGCACCTCATCTCAGAGAGTCAACAATCCTGAA 916
Db 815 TGGCTCAATGAGCAAAATTTATTTGAACTTTCTTATACCAAGACCTTCAACAATCCAGAT 874
Qy 917 GATGACAAAGTATCTTTTCTTCCGTGAAATGCAATAGATGAGAACTCTGGAAAA 976
Db 875 GATGATAAATATATTTCTTCTTCTGTAATCATCTCAAGAGGCGAGTACCTCCGATAA 934
Qy 977 GCTACTCAGCTAGAAATAGTCAGATATGCAAGATGACTTTGGAGGGCAGCAAGATCTG 1036
Db 935 ACCATCTTTCTCGAGTTTGAAGAGTTTGTAAAGATGATGAGGAGCAAGCGACGCTG 994
Qy 1037 GTGAATAAATGACACATCTCTCAAGCTCTGATTTGCTCAGTGCAGGTCCAAAT 1096
Db 995 ATAACAAGTGGACGACTTTCTTAAGGCCAGACTGATTTGCTCAATCTCTGGAAGTGAT 1054
Qy 1097 GGCATTGACACTCATTTTGAATGAACTGCGAGATGATCTCAATGAATCTTAAAGATCCT 1156
Db 1055 GGGCAGATACTTCTTGAAGCTTCAAGATATTTATTTACTCCCAAGAGATGAA 1114
Qy 1157 AAAAAATCCAGTTGATATGAGGTTTACGACTTCCAGTAAATTTTCAAGGATCAGCC 1216
Db 1115 AGAAATCCCTGTAGTATGAGCTTTTACTACAACAGCTCCATCTTCAAGGCTCTGCT 1174
Qy 1217 GTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1276
Db 1175 GTTTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1234
Qy 1277 GATGACCAATATCAATGAGTGTATCAAGGAGAGTCCCTATCCAGGCGAGCA 1336
Db 1235 GAAAGTGACAGCATCTGTTGGGTGAGTATGATGAGGAAATCTTATCCAGGCTGCT 1294
Qy 1337 ACTGTGCCAGCAAAACATTTG---GTGGTTTTGACTTCAAAAGGACCTTCTCGATGAT 1393
Db 1295 ACATGTCCAAGCAAAACCTATGACCACTGATTAAGTCCACCGAGATTTTCCAGATGAT 1354
Qy 1394 GTTATAACCTTTGCAAGAGTATCCAGCCATGTAATCCAGTGTTCCTATGAACAT 1453
Db 1355 GTCATCAGTTTCATAAAGCGGCACTGTGTATGATTAAGTCCGATATACCAGTTCCAGGA 1414
Qy 1454 CGCCCAATAGTATCAAAAGGATGTAATTTATCAATTTACAAATCTGCTAGACCGA 1513
Db 1415 GGACCAAGTTCAAGAGATCAATGTGGATTAAGTCAAGTCAAGATGATGATGATGATGAT 1474
Qy 1514 GTGGATGCAAGAGATGAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1573
Db 1475 GTCAATGCAAGAGATGGCCAGTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1534
Qy 1574 CTTAAAGTGTCTCAATCTCAAGAGACTTGGTATGATTTAGAGAGGTTCTGCTGGA 1633
Db 1535 CTCAAAGTGTCTCAGCATTTTCAAGGAAAGAGTGG---AATATGGAAGAGGAGTGTCTGGAG 1591
Qy 1634 GAAATGACAGTTTCTTGGGAAACCGACTGCTATTTTCAAGCAATGAGCTTCCACTAAGCAG 1693
Db 1592 GAGTTGAGATATTTCAAGCACTCATCAATCATCTTGAACATGGAATGTCTCTGAAGCAG 1651
Qy 1694 CAACAATATATATTTGGTTTCAACGGCTGGGGTTGGCCAGCTCCCTTTTACACCGGTGTGAT 1753
Db 1652 CAACAATGTATGATTTTCCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1711
Qy 1754 ATTTACGGAAAGCGTGTGCTGAGTGTGCTGCTGCGCGAGACCTTTACTGTCTTGGAT 1813
Db 1712 ACTTATGGGAAAGCTTGGCAGACTGTTGTCTTGGCAGAGACCTCTACTGTGCTCTGGAT 1771
Qy 1814 GTTCTGCAATGTTCTGCTATTTTCCACTGCAAGAGAGCAGCAAGAGCAAGATATA 1873
Db 1772 GGAATGCAATGCTCTCGATATGCTCTCTTCTTCTTAAAGGAGAGCTAGAGCCCAAGATGTA 1831

RESULT 12

US-10-176-482-347

; Sequence 347, Application US/10176482

; Publication No. US20030022296A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: ACIDS ENCODING THE SAME

; CURRENT APPLICATION NUMBER: US/10/176,482

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 347

; LENGTH: 3871

; TYPE: DNA

; ORGANISM: Homo Sapien

US-10-176-482-347

Query Match

Best Local Similarity 26.5%; Score 718.6; DB 14; Length 3871;

Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;

Qy 269 AATGGGAAGAACAAATGTGCCAAGGCTGAAATTTATCTTACAAAGAAATGTTGGAATCCAAC 328
Db 215 ACTTTGAAGAAATAATTTCCAAAGACTCAAGTAACTTACAAAGACTTCTGCTTTCAAT 274
Qy 329 AATGTGATCACTTCAATGGCTTGGCCCAACAGCTCCAGTTATCATACCTTCTTTTGGAT 388
Db 275 AGCTGATATCCCTTTTGGGGTTTCATCAGAGGACTGGATTTTCAAACTCTTCTCTAGAT 334
Qy 389 GAGGACGGATGAGCTGTATGTTGGAGCAAGGATCATATTTTCTTATTCGACTGTT 448

Db 335 GAGGAAAGAGGAGGCTGCTCTTTGGGAGCAAGAACACATCTTTCTCACTGCTGGTT 394
Qy 449 AA---TATCAAGGATTTTCAAGAGATTGTGCGCCAGTATCTTACACAGAGAGATGAA 505
Db 395 GACTTAAACAAAATTTTAAAGAGATTATTTGGCCCTGCTGCAAGGAAAGGGTGAATTA 454
Qy 506 TGAAGTGGGCTGGAAAGACATCTCTGAAAGAAATGTGCTAAATTTTCAATCAAGGTACTTTAAG 565
Db 455 TGTAAATAGCTGGGAAAGATGCCAATACAGATGTGCAATTTTCAATCAGAGTACTTCAG 514
Qy 566 GCATATAATCAGACTCACTGTGACGCTGTGGAACGGGGGCTTTTCAATCAATTTTGCACC 625
Db 515 CCCTATAACAAAACCTCACATATATGTGTGGAACCTGGAGCATTTTCAATCAATATGTGG 574
Qy 626 TACATTTGAATTTGGACATCATCTGAGGACAAATATTTTAAAGCTGGAGAACTCACATTTT 685
Db 575 TATATTTGATCTTGGAGTCTCAAGGAGGATATTTATTTCAAACTAGACACACATATTTG 634
Qy 686 GAAACGGCGGTGGGAAGAGTCCATATGACCCCTGCTGACAGCATCCCTTTTAAATA 745
Db 635 GAGTCTGGCAGACTGAAATGTCTTTTCGATCTCAGCAGCCTTTTGTCTCAGTAATGACA 694
Qy 746 GATGGAGAAATTAATCTCTGAACTGACGTGATTTTATGGGCGAGACTTTGCTATCTTC 805
Db 695 GATGAGTACCTCTACTCTGGAACAGCTTCTGATTTCTTTGGCAAGATATCTGCAITCACT 754
Qy 806 CGAATCTTTGGG- - - - - ACCACCAACCAATCAGGACAGGACGATGNTCCAGG 856
Db 755 CGATCCCTTTGGGCTTACTCATGACCACTTACATCAGAACTGACATTTTCAAGGACACTAC 814
Qy 857 TGCTCAATGATCCAAAGTTTCAATAGTGCCCACTCATCTCAGAGAGTGCACAACTCTGAA 916
Db 815 TGCTCAATGGAGCAAAATTTATTGGAACTTTCTTATACACAGACACTTCAATCCAGAT 874
Qy 917 GATGACAAAGTATATCTTTTCTTCCGTGAAATGCAATAGATGGAGAACTCTGGAATA 976
Db 875 GATGATAAAATATATTTCTTTCTGTAATCATCTCAAGAAAGGAGTACCTCCGATAAA 934
Qy 977 GCTACTCACGCTAGATAGTGCAGATATGCAGAAATGACTTTGGAGGACAGAGATCTG 1036
Db 935 ACCATCTTTCTGAGTGGAGAGTTTGTAAAGATGATGTAGGAGGACAACGCGACCTG 994
Qy 1037 GTGAATAAATGGACAACTTCTCAAGCTCGTCTGATTTGCTCAGTGCAGGTCGCAAT 1096
Db 995 ATAAACAAGTGGAGCACTTTCTTAAGGCGAGACTGATTTGCTCAATCTCTGGAAGTAT 1054
Qy 1097 GGCATGACACTCATTTTGAAGAACTGCGAGGATGATTTCTTAATGAATTTTAAAGATCT 1156
Db 1055 GGGCGAGATACTTACTTTGATGAGCTTCAAGATATTTTACTTCCCAACAAGAGATGAA 1114
Qy 1157 AAAAATCCAGTTGATATAGGAGTTTACGACTTCCAGTAAATTTTCAAGGATCAGCC 1216
Db 1115 AGAAATCTGTAGTATGAGGCTTTTACTAACCAAGCTTCAATCTTCAAGGCTCTGCT 1174
Qy 1217 GTGTGATGATATAGCATGATGTGAGAAGGCTTCTTGGTCCATATGCCACAGG 1276
Db 1175 GTTTGTGTATAGCATGCTGCATCAGAGCACTTTTAAAGTCCATATGCTATAAG 1234
Qy 1277 GATGGACCAACTATCAATGGGTGCTTATCAAGGAAAGATGCCCTATCCAGCGGACGA 1336
Db 1235 GAAAGTGCAGACCACTGTTGGGTGAGTATGATGGGAGAAATCTTATCCAGGGCTGTG 1294
Qy 1337 ACTGTGCCAGCAAAATTTG- - - GTGGTTTGTACTTCAAGAGACCTTCTGATGAT 1393
Db 1295 ACATGTCCAAAGCAAAACCTTATGACCACTGATTAAGTCCACCGAGATTTTCCAGATGAT 1354
Qy 1394 GTTATAACCTTTTCAAGAAAGTATCCAGCCATGTACAATCCAGTGTTCCTATGAACAT 1453
Db 1355 GTCATCAGTTTCAATAAGCGGCACTCTGTGATGTATAAGTCCGTATACCCAGTTGCAGA 1414
Qy 1454 CGCCCAATAGTATCAAAACGGAGTAAATTTTCAATTTACAAATTTGCTGTAGACGA 1513
Db 1415 GGACCAACGTTCAAGAGAAATCAATGTGGATTACAGACTGACAGATAGTGTGGATCAT 1474

Qy 1514 GTGGATGAGAGAGTGGACAGTATGATGTTATTTTATCGGAAACAGATGTTTGGGACCGTT 1573
Db 1475 GTCTTGGAGAGAGTGGCAGTACGATGATATTTTCTTGGNAACAGACATTTGGNACTGTC 1534
Qy 1574 CTTAAAGTATGTTCAATTTCTAAGGAGACTTGGTATGATTTTGAAGAGGTTCTGTGGA 1633
Db 1535 CTCAAAGTTGTGAGCATTTTCAAGAGAAAGTGG- - - AATATGGAAGAGGTAGTGTCTGGAG 1591
Qy 1634 GAATGACAGTTTTCGGGAACCGACTGCTATTTTCAAGCAATGGAGCTTTCCACTAAGCAG 1693
Db 1592 GAGTTTGCAGATATTTCAAGCACTCATCAATCTTTTGAACATGGAATTTGTCTCTGAAGCAG 1651
Qy 1694 CAACAATATATATTTTGGTTTCAACGGCTGGGGTTGGCCAGCTCCCTTTTACACCGGTGTGAT 1753
Db 1652 CAACAAATTTGATATTTGGTTCCGAGATGATTTAGTTTCACTCTCTTGCACAGATGCGAC 1711
Qy 1754 ATTTACGGGAAGCGTGTGCTGAGTGTTCCTGCGCCGAGACCTTTACTGTGCTTGGAT 1813
Db 1712 ACTTATGGGAAGCTTTGGCAGACTGTTGTCTTGCAGAGACCCCTACTGTGCTCTGGAT 1771
Qy 1814 GGTTCGATGTTCTGCTATTTTCCCACTGCAAGAGACGCAAGAGCAGACAGATATA 1873
Db 1772 GGAATGATGCTCTCTCGATATGCTCTTCTTAAAGGAGAGCTAGACGCCAAGATGTA 1831
Qy 1874 AGAATGAGAGACCACTGACTCAGTCTTTCAGACTTTTACACCATGATAATCACCATGCCAC 1933
Db 1832 AATATGCGGACCCATCACCAGTGTGGAGCATCGAAGACAGCATTTAGTCAATG- - - NA 1888
Qy 1934 AGCCCTGAAGAGAGATCATCTATGTTGTAGAGATAGTAGACATTTTGGAAATGCAAT 1993
Db 1889 ACTGCTGATGAAAGGTGATTTTGGCATTTAACTCAACTCAACTTTCTGGAATGTATA 1948
Qy 1994 CGAAGTCCGAGAGAGCGCTGCTATTTGGCAATTCAGAGCGGCAATTAAGAGCGGAAA 2053
Db 1949 CCTAAATCCCAACAGCAACTATTAAATGGTATATCCAGAGGTCAAGGAGTGAATCGA 2008
Qy 2054 GAAGAGATCAGATGGATGATCATCATCAGGACAGATCAAGGCTTCTGCTAGCTAGT 2113
Db 2009 GAGGATTTGAAGCCGATGAAGATCATCAACGGATATTTGGCTACTGATTCGAGT 2068
Qy 2114 CTACAACAGAGATTTAGGCAATTTACCTTGCATCGGTTGGAAACATGTTTATACAA 2173
Db 2069 TTGCAGAGAGAGATTTCTGGATGATTTACTGCAAGGCCAGGAGCACATTTTCATCCAC 2128
Qy 2174 ACTCTTTTAAAGTAAACCTTGAAGTCAATTCACACAGAGCATTTTGGAGAACTTTCTCAT 2233
Db 2129 ACCATAGTGAAGTGAATTTGAATGTCATTGAGAAATGAACAGATGGAATAATCCAGAGG 2188
Qy 2234 AAAGATGATGATG 2246
Db 2189 GCAGAGCATGAGG 2201

RESULT 13

US-10-176-757-347

; Sequence 347, Application US/10176757

; Publication No. US20030022297A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Deanovers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; TITLE OF INVENTION: ACIDS ENCODING THE SAME

; FILE REFERENCE: P3430R1C86

```
; CURRENT APPLICATION NUMBER: US/10/176,757
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 347
; LENGTH: 3871
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-757-347

Query Match          26.5%; Score 718.6; DB 14; Length 3871;
Best Local Similarity 61.9%; Pred. No. 7,9e+205;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;

QY 269 AATGGGAAGAACAAATGTGCAAGGCTGAAATTTATCCTCAAAAGAAATGTGGAATCCAAC 328
DB 1157 AAAAAATCCAGTTGTATATGAGAGTGTTCACGACTTCCAGTAACATTTTCAAGGGATCAGCC 1216
QY 215 ACTTTGAAGCAAAATATTCGAAGACTCAAGCTAACCTACAAAGACTTGTCTTCAAT 274
DB 1115 AGAAATCTGTAGTATATGAGAGTCTTTACTACACACGACTCCATCTTCAAGGCTCTGCT 1174
QY 329 AATGTGATCACTTTCAATGGCTTGGCCAAAGCTCAGTATATCACTTCTTTTGGAT 388
DB 1217 GTGTGTATGTATAGCATGAGTGTGAGAGGGTGTTCCTTGGTCCATATGATGCCACAGG 1276
QY 275 AGCTGTATTTCCCTTTTGGGTTTCATCAGAAGGACTGGATTTTCAAACTCTTCTTAGAT 334
DB 1175 GTTTGTGTATAGCATGGCTGACATCAGAGCAGTTTTTAAATGGTCCATATGCTCATAG 1234
QY 389 GAGGAACGGAGTAGGCTGTATGTTGGAGCAAGAGATCACATATTTTCATTCGACCTGTT 448
DB 1277 GATGACCCAACTATCAATGGTGTCTTATCAAGGAAGAGTCCCTCTATCCAGGCCAGGA 1336
QY 335 GAGGAAGAGCGAGGCTGCTCTTGGAGCCAAAGACCAATCTTTCTACTAGTCTGGTT 394
DB 1235 GAAAGTCAGACACCATCTGTTGGTGCAGTATGATGGGAGAAATTCCTTATCCACGGCTGGT 1294
QY 449 AA---TATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACACAGAGAGATGAA 505
DB 1337 ACTTGTCCAGCAAAACAATTTG---GTGGTTTTGTACTCTACAAAGAGACTTCTCTGATGAT 1393
QY 395 GACTTAAACAAAATTTTAAAGAGATTTATTTGGCCTGTGCTGCAAGGAGGGTGAATTA 454
DB 1295 ACATGTCCAAAGCAAAACCTATGACCCACTGATTAAGTCCACCGAGATTTTCCAGATGAT 1354
QY 506 TGCAGTGGCTGGAAAGACATCTCTGAAAGAAATGTGCTAAATTTTCATCAAGGTACTTAAG 565
DB 1394 GTTATAACTTTTGCNAAGAGTCAATCCAGCCATGATCAATCCAGTGTTCCTTATGAACAT 1453
QY 455 TGTAAATAGCTGGGAAAGATGCCAATACAGAATGTGCAAAATTTTCATCAGAGTACTCAG 514
DB 1355 GTCATCAGTTTTATTAAGCGGCACCTCTGTGATGATATAAGTCCGTATATCCCAAGTTGACGA 1414
QY 566 GCATATATCAGACTCAGCTTGTACCTGTGGAACGGGGCTTTTCATCCAAATTTGCACC 625
DB 1454 CGCCCAATAGTATGATCAAAACCGATGATAATTTATCAAAATTTTACAAATTTGTCTAGACGA 1513
QY 515 CCTATATAAAACCTACATATATGTGTGGAATGGAAGCAATTTTCATCCAAATATGTGG 574
DB 1415 GGAACCAAGCTTCAAGAGAAATCAATGTGATTTACAGCTGACACAGTATGTTGATCAT 1474
QY 626 TACATTTGAAATTTGGACATCATCTCAGGACAAATATTTTAAAGCTGGAGAACTCACATTT 685
DB 1514 GTGATGCAGAGATGGAAGTATGATTTATGTTTATCGAAACAGATGTTTGGGACCGTT 1573
QY 575 TATATTGATCTTGGAGCTTACAAGAGAGATATTTATTTCAACTAGACACATATTTTG 634
DB 1475 GTCATTTGAGAAGATGGCCAGTACGATGATGTTTCTTGGAAACAGACATTTGGAATCTGTC 1534
QY 686 GAAACGGCTGGGAAGAGTCCATATGACCTTAAGCTGTGACAGCATTCCTTTTAATA 745
DB 1574 CTTAAAGTAGTTTCAATTTCTTAAGAGACTTGGTATGATTTTAGAAGAGGTTCTTCTGTGAA 1633
QY 635 GAGTCTGGCAGACTGAAATGTCTTTCGATCTCAGCAGCCTTTTGTCTTCAATATGACA 694
DB 1535 CTCAAAGTTGTGAGCATTTTCAAGGAAAGTGG---AATATCGAAGAGGTATGTTGTTGAG 1591
QY 746 GATGAGAAATATATCTCTGGAACCTCAGTGTATTTATGGGCGAGACTTTCCTATCTTC 805
DB 1634 GAAATGACAGTTTTTCCGGAACCGACTGCTATTTTTCAGCAATGGAGCTTTCCACTAAGCAG 1693
QY 695 GATGATACCTCTCTCTGGAACAGCTTCTGATTTTCTTGGCAAGAGTACTGCACTTCACT 754
DB 1592 GAGTTGCAAGATTTCAAGCACTCATCAATCATCTTGAACATGGAATTTCTCTCTGAAGCAG 1651
QY 806 CGAACTCTTTGGGC-----ACCACCAACCAATCAGGACAGCAGCATGATTCAGG 856
DB 1694 CAACAACTATATATTTGGTTTCAACGGCTGGGGTGTGCCAGCTCCCTTTTACACGGGTGTGAT 1753
QY 755 CGATCCCTTTGGCCTTACTCATGACCAACCACTACATCAGACTGACATTTTCAGAGCACTAC 814
DB 1652 CAACAAATTTGATTTGTTTCCGAGATGGAATTTAGTTTCACTCTCTCTGACAGATGCGAC 1711
QY 857 TGGCTCAATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGAACAATCCTGAA 916
DB 1754 ATTTACGGGAAGAGGTGTGCTGAGTGTGCTCGCCGAGACCTTACTTCTGCTTGGGAT 1813
QY 815 TGGCTCAATGGAGCAAAATTTATTTGAACTTTCTTATACACAGACACCTCAATCCAGAT 874
DB 1712 ACTTATGGAAGAGCTTGGCAGACTTGTGTTGCTGCGAGAGACCTTACTTGTGCTGCTGGAT 1771
QY 917 GATGCAAGATATATCTTTTCTCGTGAATGCAATAGTAGTGGAGAACACTCTGGA 976
DB 1814 GGTTCGCTCATCTCTCGCTATTTTCCCACTGCAAGAGACGCAAGAGACGACAGATATA 1873
QY 875 GATGATAAAATATATTTCTTTCTTCTGTAATCATCTCAAGAGGCGAGTACCTCCGATAAA 934
DB 1772 GGAATGCTGCTCTCGATATGCTCTTCTTAAAGGAGAGCTTAGACGCCAAGATGTA 1831
QY 977 GCTACTCAGCTGAGATAGGTGAGATATGCAAGAAATGACTTTTGGAGGGCAGACAGTCTG 1036
DB 755 CGATCCCTTTGGCCTTACTCATGACCAACCACTACATCAGACTGACATTTTCAGAGCACTAC 814
QY 935 ACCATCTTCTCGAGTTTGGAGAGTTTGTAGAAATGATGTAGGAGGACACGAGGCTG 994
DB 857 TGGCTCAATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGAACAATCCTGAA 916
QY 1037 GTGAATAAATGACAACATCTCTCAAGCTCTGATTTGCTGATTTGCTCAGTCCAGGTTCCAAAT 1096
DB 815 TGGCTCAATGGAGCAAAATTTATTTGAACTTTCTTATACACAGACACCTCAATCCAGAT 874
QY 995 ATAAACAGTGGACGACTTTTCTTAAGGCCAGACTGATTTGCTCAATTTCTTGGAGTGTAT 1054
DB 917 GATGCAAGATATATCTTTTCTCGTGAATGCAATAGTAGTGGAGAACACTCTGGA 976
QY 1097 GGCATTGACACTCATTTTATGTAAGTCTGAGGATGATTTCTTAAATGAACTTTTAAAGATCCT 1156
DB 875 GATGATAAAATATATTTCTTTCTTCTGTAATCATCTCAAGAGGCGAGTACCTCCGATAAA 934
QY 977 GCTACTCAGCTGAGATAGGTGAGATATGCAAGAAATGACTTTTGGAGGGCAGACAGTCTG 1036
DB 935 ACCATCTTCTCGAGTTTGGAGAGTTTGTAGAAATGATGTAGGAGGACACGAGGCTG 994
QY 1037 GTGAATAAATGACAACATCTCTCAAGCTCTGATTTGCTGATTTGCTCAGTCCAGGTTCCAAAT 1096
DB 995 ATAAACAGTGGACGACTTTTCTTAAGGCCAGACTGATTTGCTCAATTTCTTGGAGTGTAT 1054
QY 1097 GGCATTGACACTCATTTTATGTAAGTCTGAGGATGATTTCTTAAATGAACTTTTAAAGATCCT 1156
DB 1055 GGGGCGAGATACTTACTTTTATGAGCTTCAAGATATTTATTTTACTCTCCCAAGAGATGAA 1114
```

QY 2234 AAAGATGATGATG 2246
DB 2189 GCAGAGCATGAGG 2201

RESULT 14

US-10-176-913-347
; Sequence 347, Application US/10176913
; Publication No. US20030022298A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C66
; CURRENT APPLICATION NUMBER: US/10/176,913
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 347
; LENGTH: 3871
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-176-913-347

Query Match 26.5%; Score 718.6; DB 14; Length 3871;
Best Local Similarity 61.9%; Pred. No. 7.9e-205;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;
QY 269 AATGGAAGAACAAATGTGCCAAGGCTGAATATCTCAAAAGAAATGTTGGAATCCAAC 328
DB 215 ACTTTGAAGCAAAATATTCAGACTCAAGCTTAACCTCAAAAGACTTGTGCTTTCAAT 274
QY 329 AATGTGATCACTTCAATGGCTTGGCCAAACAGCTCCAGTTATCAVACCTTCCCTTTGGAT 388
DB 275 AGCTGTATTCCTTTTGGGTTTCATCAGAAGGACTGGATTTTCAAACCTCTCTTAGAT 334
QY 389 GAGGAACGAGTAGGCTGTATGTTGGAGCAAAAGATCAATATTTTCATGACCTGGTT 448
DB 335 GAGGAAGAGGCGAGGCTGCTTTGGGAGCCAAAGACCACATCTTTCTACTCAGTCTGGTT 394
QY 449 AA---TATCAAGATTTTCAAAGATTTGTGCCAGTATCTTACACCAGAGAGATGAA 505
DB 395 GACTTAAACAAAATTTTAAAGATTTATTTGGCTGCTGCAAGAGGAGCGGTGGAATTA 454
QY 506 TGCAGTGGGCTGGAAGACATCTCTGAAGAAATGTGCTTAAATTTTCATCAAGTACTTAAG 565
DB 455 TGTAAATAGCTGGGAAGATGCCAATACAGATGTGCAATTTTCATCAGAGTACTTCAG 514
QY 566 GCATATAATCAGACTCACCTTGTACGCTGTGGAACGGGGCTTTTCATCAATTTGACCC 625
DB 515 CCCTATAACAAAACCTCACATATATGTGTGGAACCTGGAGCATTTTCATCCAATATGTGG 574
QY 626 TACATTGAATTTGGACATCATCTCAGGACATATTTTAAAGCTGGAGACTCACATTTT 685
DB 575 TATATTGATCTTGGAGCTTCAAGGAGGATATTTATTTCAAACCTAGACACATATTTG 634
QY 686 GAAACCGGCGGTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGCATCCCTTTTAAATA 745
DB 635 GAGTCTGGCAGACTGAAATGTCTTTTCGATCTCAGCAGCCTTTTGTCTTCAATATGACA 694
QY 746 GATGGAAATATATCTCTGGAATCGAGCTGATTTTATGGGGCGAGACTTTGCTATCTTC 805
DB 695 GATGAGTACTCTTACTCTGGAACAGCTTCTGATTTTCTTTGGCAAGAGATACCTGCTTCACT 754

QY 806 CGAACTCTTGGGC-----ACCACCAACCAATCAGGACAGAGCAGCATGATTCACGG 856
DB 755 CGATCCCTTGGGCTTACTCATGACCACCACTACATCAGAACTGACATTTTCAGAGCACTAC 814
QY 857 TGGCTCAATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGACAAATCTGAA 916
DB 815 TGGCTCAATGGAGCAAAATTTATTGGAACCTTTCTTATACACAGACACCTTACAAATCCAGAT 874
QY 917 GATGACAAAGTATATCTTTTCTCGTGAAATGCAATAGATGGAGAACACTCTGGAAA 976
DB 875 GATGATAAAATATATTTCTTCTTCTCGTGAATCATCTCAAGAGAGGAGGAGTACCTCCGATAAA 934
QY 977 GCTACTCAGCTAGAAATAGGTGAGATGCAAGATGACTTTGGAGGGGACAGAGAGTCTG 1036
DB 935 ACCATCCTTTCTCGAGTTGGAAGAGTTTGTAAAGATGATGAGGAGGACAGAGCCTG 994
QY 1037 GTGAATAAATGGACAACATCTCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCCAAAT 1096
DB 995 ATAAACAAGTGGAGACTTTTCTTAAGCCAGACTGATTTGCTCAATTTCTGGAAGTGAT 1054
QY 1097 GGCATTGACACTCATTTTGTAGAACTGCAGATGTATTTCTTAATGAACTTTAAAGATCCT 1156
DB 1055 GGGCAGATATCTTACTTTGATGAGCTTCAAGATATTTATTTACTCCCAAGAGATGAA 1114
QY 1157 AAAAATCCAGTTGTATATGAGAGTGTTCACGACTTCCAGTAACTTTCAAGGGATCAGCC 1216
DB 1115 AGAATCTCTGATGATATGAGAGTCTTTACTAACCAGCTCATCTTCAAGGCTCTGCT 1174
QY 1217 GTGTGATGATAGCATGAGTGTGAGAGGGTGTCTTGTGTCATATGCCACAGG 1276
DB 1175 GTTTGTGTATAGCATGGCTGACATCAGAGCAGTTTTTAATGTTCCATATGCTCATAAG 1234
QY 1277 GATGACCCAACTATCAATGGGTGCTTATCAAGAGAGAGTCCCTATTCACGGCCAGGA 1336
DB 1235 GAAAGTGAGACACCATCGTTGGGTGAGTATCATGGAGAAATTTCTTATTCACAGGCTGCT 1294
QY 1337 ACTTGTCCCAACAAAACATTTG---GTGGTTTTCAGCTTCAAAAGAGACTTCTCTCATGAT 1393
DB 1295 ACATGTCCAGCAAAACCTATGACCCACTGATTAAGTCCACCCGAGATTTTCCAGATGAT 1354
QY 1394 GTTATAACCTTTGCAAGAGTATCCAGCCATGATCAATCCAGTGTTCCTTATGAACAAAT 1453
DB 1355 GTCATCAGTTTCATAAAGCGGCACTCTGTGATGATTAAGTCCGTATATCCAGTTTCAGGA 1414
QY 1454 CGCCCAATAGTATCAAAACGAGTAAATTTCAATTTTACAAATTTGCTGTAGACCA 1513
DB 1415 GGACCAACGTTTCAAGAGAAATCAATGTGGATTTACAGACTGACACAGATAGTGTGATCAT 1474
QY 1514 GTGGATGCAGAAAGATGAGCAGTATGATTTATGTTTATCGGAACAGATGTTGGGACCGTT 1573
DB 1475 GTATTGCAAGAGATGGCCAGTACGATGATATGTTTCTTGGAAACAGATTTGGAATCTGTC 1534
QY 1574 CTTAAAGTAGTTTCAAATTCCTAAGGAGACTTGGTATGATTTTGAAGAGGTTCTGCTGAA 1633
DB 1535 CTCAAAGTTGTGAGCATTTTCAAAGGAAAGTGG---AATATGGAAGAGGATGATGCTGGAG 1591
QY 1634 GAAATGACAGTTTTCGGGAACCGACTGCTATTTTACGCAATGAGAGCTTTCCACTTAAGCAG 1693
DB 1592 GAGTTGACAGATTTTCAAGCACTCATCAATCATCTTGAACATGGAATTTCTCTCTGAAGCAG 1651
QY 1694 CAACAACTATATATTTGGTTTCAACGGCTGGGGTGTGCCAGCTCCCTTTTACACGGGTGAT 1753
DB 1652 CAACAAATGTACATTTGGTTCCGAGATGATTTAGTTTCACTCTCTCTTGCACAGATGCGAC 1711
QY 1754 ATTTACGGGAAGCGTGTGCTGAGTGTGCTCCCGAGAGACCTTTACTGTCTTGGGAT 1813
DB 1712 ACTTATGGGAAGCTTGGCAGACTGTGTCTTGCAGAGAGCCCTACTGTGCTCTGGAT 1771
QY 1814 GGTTCGATGCTTCTCGCTATTTTCCCACTGCAAGAGACCGCACAGACAGACAGATATA 1873
DB 1772 GGAAATGCATGCTCTCGATATGCTCTTCTTCTTAAAGGAGAGCTAGACGCCAAGATGTA 1831

QY 1874 AGAATGGAGACCCACTGACTCTCTTCTGAGCTTACACCATGATTAATCACCATGGCCAC 1933
Db 1832 AAATATGGGACCCAAATCAACCCAGTCTGGGACATCGAAGACAGCATTAAGTATGTCATG---AA 1888
QY 1934 AGCCCTGGAAGAGAGAAATCATCTATGCTGTAGAGAAATAGTAGACATATTTTGGAAATGCACT 1993
Db 1889 ACTGCTGATGAAAGGTGATTTTGGCATTTGATTTAACTCAACCTTCTGGAATGTATA 1948
QY 1994 CGAAGTCCGAGAGAGCGCTGTCTATTTGGCAATTTCCAGAGGCGAAATGAAGAGCGAA 2053
Db 1949 CCTAAATCCCAACAGCAACTATTAATGGTATATCCAGAGGTTCAGGGATGAGCATCGA 2008
QY 2054 GAAGATCAGATGATGATCATATCATCAGGACAGATCAAGGCCCTTCTGCTACGTAGT 2113
Db 2009 GAGGAGTTGAAGCCCGGAAAGAAATCATCAAAACGGAAATATGGGTACTGATTCGAAGT 2068
QY 2114 CTACACAGAGGATTCAGGCAATTAACCTCTGCGATGCGGTGGAACATGGGTTCATACAA 2173
Db 2069 TTGCAGAGAGGATTTCTGGATGTATTAATCTGCAAGCCAGGACACACTTTCATCCAC 2128
QY 2174 ACTTCTTAAAGTAACCTCGAAGTCAATTTGACACAGAGCATTTTGGAAAGCACTTCTTCAT 2233
Db 2129 ACCATAGTGAAGTCACTTTGAATGTCAATTGAGATGAACAGATGGAATAATCCAGAGG 2188
QY 2234 AAGATGATGATG 2246
Db 2189 GCAGAGCATGAGG 2201

RESULT 15

US-10-180-552-347
; Sequence 347, Application US/10180552
; Publication No. US20030022300A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Destoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C153
; CURRENT APPLICATION NUMBER: US/10/180,552
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 347
; LENGTH: 3871
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-180-552-347

Query Match 26.5%; Score 718.6; DB 14; Length 3871;
Best Local Similarity 61.9%; Pred. No. 7.9e-205;
Matches 1233; Conservative 0; Mismatches 739; Indels 21; Gaps 5;
QY 269 AATGGGAAGAACAAATGTGCCAAGGCTGAAATTTATCCTACAAAGAAATGTGGAATCCAAC 328
Db 215 ACTTTGAAGCAAAATATTCAGAGCTCAAGCTTAACCTACAAAGACTTGTCTGCTTTCAAT 274
QY 329 AATGTGATCACTTCAATGGCTTGGCCAAACAGCTCAGTTATCAATCCTTCTTTTGGAT 388
Db 275 AGCTGTATTCCTTTTGGGTTTCATCAGAAAGGACTGGATTTTCAAACTCTTCTCTTAGAT 334
QY 389 GAGGAACGAGTAGGCTGTATCTTGGAGCAAGAGATCATATTTTTCATTCAGCTGGTT 448
Db 335 GAGGAAGAGGAGGCTGTCTTTGGAGGCCAAAGACCATCTTCTACTAGTCTGGTT 394

QY 449 AA---TATCAAGGATTTTCAAAGAAATGTGTGGCCAGTATCTTACACCAGAAAGATGAA 505
Db 395 GACTTAAACAAAAATTTTAAAGAAATTTATTTGGCCTGCTGCAAGGAACCGGTGAATTA 454
QY 506 TCCAGTGGCTGGAAAGACATCTCTGAAGAAATGTCTTAATTTTTCATCAAGGTACTTAAG 565
Db 455 TGTAATTTAGCTGGGAAAGATGCCAATACAGAAATGTCAAAATTTTTCATCAGAGTACTTCAG 514
QY 566 GCATATAATCAGACTCACTTGTACCGCTGTGGAACGGGGCTTTTTCATCCAATTTTGCACC 625
Db 515 CCTATAACAAACTCACAATATATGTGTGGAACCTGGAGCAATTTTCAATCATATGTGG 574
QY 626 TACATTTAAATTTGGACATCATCTCTGAGCAATATTTTAAAGCTGGAGAACTCAATTTT 685
Db 575 TATATTTGATCTTGGAGTCTACAAGGAGGATATTAATTTCAAACCTAGACACACATAATTG 634
QY 686 GAAACGGCCGTGGGAAAGTCCATATGACCTTAAGCTGCTGACAGCATCCCTTTTAATA 745
Db 635 GAGTCTGGCAGACTGAAATGTCTTTCATCTCAGCAGCCTTTTGTCTTCAGTAATGACA 694
QY 746 GATGAGAAATATATCTCTGGAACCTGAGCTGATTTTATGGGCGAGACTTTGCTATCTTC 805
Db 695 GATGAGTACCTTCTCTGGAACAGCTTCTGATTTCTTGGCAAGAACTATGCAATTCAT 754
QY 806 CGAATCTTTGGGC-----ACCACCAACCAATCAGGACAGACAGCATGATTTCCAGG 856
Db 755 CGATCCCTTGGCCCTTACTCATGACCACCACTACATCAGAACTGACATTTTCAGAGCACTAC 814
QY 857 TGGCTCAATGATCCAAAGTTCAATTAAGTCCCACTCATCTCAGAGAGTACATCTGNA 916
Db 815 TGGCTCAATGAGGCAAAATTTTATGGAACCTTTCTTCATACCAAGACCTTACATCAAGAT 874
QY 917 GATGACAAAGTATATCTTTTCTCGTGAATAATGCAATAGATGGAGAACACTCTCGAAA 976
Db 875 GATGATAAAATATATTTCTTTCTGTAATCATCTCAAGAGGAGTACCTTCCATATA 934
QY 977 GCTACTCAGCTAGAAATAGGTGAGATATGCAAGAAATGACTTTTGGAGGACAGAAAGTCTG 1036
Db 935 ACCATCTTTCTCGAGTTTGGAAAGTGTGTAAGAAATGATGTAGGAGGACAAACGACGCTG 994
QY 1037 GTGAATAATGGACAAACATTTCTCAAGCTGCTGATTTGCTCAGTCCAGTCCCAAT 1096
Db 995 ATAAACAAAGTGGACGACTTTTCTTAAGGCCAGACTGATTTGCTCAATTTCTTGGAAAGT 1054
QY 1097 GCGATTGACACTCATTTTGTGAACCTGAGAGTATTTCTTAATCAACTTTTAAAGATCCT 1156
Db 1055 GGGCAGATATCTTACTTTGATGAGCTTCAAGATATTTTACTCCCAAGAGATGNA 1114
QY 1157 AAAAATCCAGTTGTATATGAGTGTGTTACGACTTTCAGTAAACATTTTCAAGGATCAGCC 1216
Db 1115 AGAAATCTGTAGTATATGAGTCTTTTACTACAAACGAGCTCCATCTTCAAGGCTCTGCT 1174
QY 1217 GTGTGTATGATGATGATGATGTGAGAGGGTGTCTTGTGCTCATATGCCCACAGG 1276
Db 1175 GTTTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1234
QY 1277 GATGACCCCAACTATCAATGGTGTCTTATCAAGGAAGTCCCTTATCCAGGCGCAGA 1336
Db 1235 GAAAGTGAGACCACTGTGTGGTGGAGTATGATGGAGAAATTCCTTATTCACGGCTGTG 1294
QY 1337 ACTTGTCCAGCAAAACATTTG---GTGGTTTGTGACTCTACAAAGGACCTTCTCTGATGAT 1393
Db 1295 ACATGTCCAAGCAAAACCTATGACCCACTGATTAAGTCCACCCGAGATTTTCCAGATGAT 1354
QY 1394 GTTATAACCTTTGCAAGAGTATCCAGCAATGATCAATTCAGTGTGTCTTATGAAACAT 1453
Db 1355 GTCATCAGTTTCAATAAAGCGGCACTCTGTGATGTATAAGTCCGTATATCCAGTGTGAGGA 1414
QY 1454 CGCCCAATAGTATCAAAACGGATGAAATTTTCAATTTTACACAAATTTGCTGTAGACGA 1513
Db 1415 GGACCAAGCTTCAAGAGAAATCAATGTGGATTACAGACTGACACAGATAGTGTGGATCAT 1474

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 09:54:25 ; Search time 27.5 Seconds
(without alignments)
9265.116 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 4896

Sequence: 1 aatctttatttatcatgatg.....aggctttttttcttaataacc 2709

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match' 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-O=/cgn2 1/USPTO spool/US09774490/runat 03082003 095416 25322/app query fasta_1.2887
-DB=Swissprot 41 -QFMT=fastan -SUFFIX=isp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MAIRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09774490 @CGN 1 1 33 @runat 03082003 095416 25322 -NCPU=6 -ICPU=3
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4201	85.8	771	1 SM3A_HUMAN	Q14563 homo sapien
2	4057.5	82.9	772	1 SM3A_RAT	Q63548 rattus norv
3	4035.5	82.4	772	1 SM3A_MOUSE	Q08665 mus musculu
4	3784.5	77.3	772	1 SM3A_CHICK	Q90607 gallus gall
5	3207.5	65.5	778	1 S21B_BRARE	Q9W686 brachydanio
6	3075.5	62.8	860	1 S21A_BRARE	Q9W7J1 brachydanio
7	2247.5	45.9	777	1 SM3D_HUMAN	Q95025 mus sapien
8	2245	45.9	761	1 SM3D_CHICK	Q90663 gallus gall
9	2125	43.4	764	1 SM22_BRARE	Q9W696 brachydanio
10	2118	43.3	749	1 SM3B_HUMAN	Q13214 homo sapien
11	2059.5	42.1	748	1 SM3B_MOUSE	Q62177 mus musculu
12	2044	41.7	775	1 SM3E_HUMAN	O15041 homo sapien
13	2020	41.3	775	1 SM3E_MOUSE	P70275 mus musculu
14	1993	40.7	785	1 SM3E_CHICK	O42377 gallus gall
15	1940	39.6	785	1 SM3F_HUMAN	Q13275 homo sapien
16	1928	39.4	785	1 SM3F_MOUSE	O88632 mus musculu
17	1894	38.7	751	1 SM3C_CHICK	O42236 gallus gall
18	1850.5	37.8	751	1 SM3C_MOUSE	Q62181 mus musculu

19	1849.5	37.8	751	1 SM3C_HUMAN	Q99985 homo sapien
20	1003	20.5	861	1 SM4D_MOUSE	Q09126 mus musculu
21	981.5	20.0	862	1 SM4D_HUMAN	Q92854 homo sapien
22	966	19.7	834	1 SM4C_MOUSE	Q64151 mus musculu
23	959.5	19.6	837	1 SM4G_MOUSE	Q9WJH7 mus musculu
24	937	19.1	838	1 SM4G_HUMAN	Q9Ntn9 homo sapien
25	889	18.2	832	1 SM4B_HUMAN	Q9Npr2 homo sapien
26	857.5	17.5	782	1 SM4B_MOUSE	Q62179 mus musculu
27	852.5	17.4	1030	1 SM6A_HUMAN	Q9H2e6 homo sapien
28	834.5	17.0	888	1 SM6A_MOUSE	Q35464 mus musculu
29	831	17.0	766	1 SM27_BRARE	Q9Ybx4 brachydanio
30	825.5	16.9	761	1 SM4A_HUMAN	Q9H3r1 homo sapien
31	802	16.4	760	1 SM4A_MOUSE	Q62178 mus musculu
32	792	16.2	771	1 SM1A_DROME	Q24322 drosophila
33	783.5	16.0	697	1 SM2A_SCHGR	Q9XZC8 schistocerc
34	780.5	15.9	706	1 SM2A_DROME	Q24323 drosophila
35	766.5	15.7	730	1 SM1A_SCHAM	Q26473 schistocerc
36	766	15.6	1093	1 SM5B_MOUSE	Q60519 mus musculu
37	760.5	15.5	930	1 SM6C_HUMAN	Q9H3t2 homo sapien
38	756.5	15.5	777	1 SM4F_MOUSE	Q92123 mus musculu
39	752.5	15.4	776	1 SM4F_RAT	Q92143 rattus norv
40	745.5	15.2	960	1 SM6C_RAT	Q9WtL3 rattus norv
41	742	15.2	931	1 SM6C_MOUSE	Q9Wtm3 mus musculu
42	737.5	15.1	770	1 SM4F_HUMAN	Q95754 homo sapien
43	737.5	15.1	1077	1 SM5A_MOUSE	Q62217 mus musculu
44	730	14.9	1074	1 SM5A_HUMAN	Q13591 homo sapien
45	723	14.8	887	1 SM6B_RAT	Q70141 rattus norv

ALIGNMENTS

RESULT 1
SM3A_HUMAN
ID SM3A_HUMAN STANDARD; PRT; 771 AA.
AC Q14563;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3A precursor (Semaphorin III) (Sema III).
GN SEMA3A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=94094332; PubMed=8269517;
RA Kolodkin A.L., Matthes D.J., Goodman C.S.;
RT "The semaphorin genes encode a family of transmembrane and secreted
RT growth cone guidance molecules.";
RL Cell 75:1389-1399(1993).
RN [2]
RP SEQUENCE OF 1-37 FROM N.A.
RA Woessner J., Minx P., Hinds K., Strommatt C.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 39-182 FROM N.A.
RA Rohlfing T., Tin-Wollam A.M., Duckels G.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
CC CONES. COULD SERVE AS A LIGAND THAT GUIDES SPECIFIC GROWTH CONES
CC BY A MOTILITY-INHIBITING MECHANISM. BINDS TO THE COMPLEX
CC NEUROPILIN-1/PLEXIN-1 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted (by similarity).
CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
CC THIRD OF THE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; L26081; AAAG5938.1; -;
 DR EMBL; AC004451; -; NOT ANNOTATED_CDS.
 DR EMBL; AC004848; AAC78622.1; -;
 DR PIR; D49423; D49423.
 DR Genew; HGNC:10723; SEMA3A.
 DR MIM; 603961; -;
 DR GO; GO:0005576; C:extracellular; TAS.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS00835; IG_LIKE; 1.
 DR Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 771 SEMAPHORIN 3A.
 FT DOMAIN 240 538 Sema.
 FT DOMAIN 580 664 IG-LIKE C2-TYPE.
 FT DOMAIN 727 769 ARG/LYS-RICH (BASIC).
 FT DISULFID 649 722 BY SIMILARITY.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 590 590 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 771 AA; 88889 MW; 9985F8D3EAD8456 CRC64;

Alignment Scores:

Pred. No.: 0 Length: 771
 Score: 4201.00 Matches: 771
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 85.80% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3A_HUMAN (1-771)

QY 200 ATGGGCTGTTAACTAGGATTCTGTCTTTCTGGGAGTATTACTTACAGCAAGACGA 259
 DB 1 MetGlyTTPLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
 QY 260 AACTATCAGAAATGGGAAGAACAAATGTGCCAGGCTGAAATATCTTACAAAGAAATGTTG 319
 DB 21 AsnTyrGlnAsnGlyLysAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
 QY 320 GAATCCACAAATGTGATCACTTCAATGCTTGGCCACAGCTCCAGTTATCATCCCTTC 379
 DB 41 GluSerAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
 QY 380 CTTTGGATGAGCAACGGAGTAGGCTGTATGTTGGAGCAAGGATCACATATTTTCATTC 439
 DB 61 LeuLeuAspGluLysArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
 QY 440 GACCTGGTTAATATCAAGATTTTCAAAGATTTGTGGCCAGTATCTTACACCAAGAGA 499
 DB 81 AspLeuValAsnIleLysAspPheGlnLysIleValTTPProValSerTyrThrArgArg 100
 QY 500 GATGATGCACTGGCTGGAAAGACATCTCTGAAGATGTGCTAAATTTTCATCAAGGTA 559
 DB 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
 QY 560 CTTAAGGCATATAAATCAGACTCACTTGTACGCTGTGGAAACGGGGCTTTTCATCCAATT 619

Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
 QY 620 TGCACCTACATGAAATTTGGACATCATCTCTAGGACAAATATTTTAACTGAGACTCA 679
 DB 141 CysThrTyrIleGluIleGlyHisProGluAspAsnIlePheLysLeuGluAsnSer 160
 QY 680 CATTTTGAAAACGGCGTGGGAAGAGTCCATATGACCTTAAGCTCTGACACATCCCTT 739
 DB 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
 QY 740 TTAATAGATGGAGAAATATATCTCTGGAACCTCAGCTGATTTTATGGGCGAGACTTGT 799
 DB 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
 QY 800 ATCTTCCGAACCTCTTGGCACCACCACCATCAGGACAGACAGACATGATTCAGGTTGG 859
 DB 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
 QY 860 CTCAATGATCCAAAGTTTCATTAGTGCACCTCATCTCAGAGAGTGACAATCCTCAAGAT 919
 DB 221 LeuAsnAspProLysPheIleSerAlaHisLeuLeuSerGluSerAspAsnProGluAsp 240
 QY 920 GACAAAGTATATCTTTTCTCCGTCAAATGCAATAGATGGAGAACACTCTGGAAGAGCT 979
 DB 241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
 QY 980 ACTCAGCTAGATAGTTCAGATATGCAAGATGACTTTGGAGGCGCAGAGCTCTGTTG 1039
 DB 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
 QY 1040 AATAAATGGCAACATCTCTCAAAGCTCGTCTGATTTGCTCAGTGCCAGGTCCAAATGGC 1099
 DB 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuLeuCysSerValProGlyProAsnGly 300
 QY 1100 ATTGACACTCAATTTGATGAACACTGCAGGATGATTTCTTAATCAACTTTAAAGATCCTAAA 1159
 DB 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
 QY 1160 AATCCAGTTGATATGGAGTGTTCAGACTTCCAGTAAACATTTTCAAGGGATCAGCCGTG 1219
 DB 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
 QY 1220 TGTATGTATAGCATGATGTGAGAGGGTGTTCCTTGTGTCCATATATGCCACACAGGAT 1279
 DB 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
 QY 1280 GGACCCCACTATCAATGGTGCCTTATCAAGGAAGAGTCCCTATCCACGCCAGGAAT 1339
 DB 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
 QY 1340 TGTCCAGCAAAACATTTGGTGGTTTGGCTTCTACAAAGGACCTTCTCTGATGATGTTATA 1399
 DB 381 CysProSerLysThrPheGlyPheAspSerThrLysAspLeuProAspAspValIle 400
 QY 1400 ACCTTTGCAAGAAGTTCATCCAGCCATGACAAATCCAGTGTTCCTATCAACAAATCCCCA 1459
 DB 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
 QY 1460 ATAGTGATCAAAACGGATGTAATTTATCAATTTTACAAATTTGCTGAGACCGAGTGGAT 1519
 DB 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
 QY 1520 GCAGAGATGACAGATGATGTTTATGTTTATCGAACAGATGTTGGACCGTCTTTAAA 1579
 DB 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
 QY 1580 GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAAGAGGTTCTCTCTGGAAGAAATG 1639
 DB 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluGluMet 480
 QY 1640 ACAGTTTTCGGGAACCGACTGCTTATTTACAAATGAGGAGCTTTCACATAGCAGCAACA 1699
 DB 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500

QY 1700 CTATATATTGTTCAACGGCTGGGTTGCCAGCTCCCTTTTACACCGGTGTGATATTAC 1759
 Db LeuTyriIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
 QY 1760 GCGAAGCGTGTGCTGAGTGTTCCTCGCCGCGAGACCCCTTACTGCTTGGATGGTCT 1819
 Db GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
 QY 1820 GCATGTTCTCGCTATTTCCTCACTGCAAGAGACGCAAGACGCAAGATATAAGAAAT 1879
 Db AlaCysSerArgTyrPheProThrAlaLysAArgThrArgGlnAspIleArgAsn 560
 QY 1880 GGAGACCCACTGACTCAGCTTTCAGACTTACACCATGATATACATGCGCCACGCCCT 1939
 Db GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisGlyHisSerPro 580
 QY 1940 GAAGAGAGAAATCATCTATGCTGTAGAGATAGTACACATTTTTCGAATGCGAGTCCGAG 1999
 Db GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
 QY 2000 TCGCAGAGACGCTGCTTATTCGCAATTCAGAGCGCAATGAAAGAGCAAGAAAGAG 2059
 Db SerGlnArgAlaLeuValTyrTrpGlnPheGlnAArgAsnGluArgLysGluGlu 620
 QY 2060 ATCAGAGTGATCATATCATATCATCAGGACAGATCAAGGCTTCTGCTAGTGTCTACAA 2119
 Db IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
 QY 2120 CAGAAGGATTACAGCAATACCTCTGCCATGGGTGGACATGCGGTTCATACAACTCTT 2179
 Db GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
 QY 2180 CTTAAGTGAACCTCGAAGTCAATTCACACAGACGATTTGGAAGAACTTCTTCATAAGAT 2239
 Db LeuLysValThrLeuGluValIleAspThrGluHisLeuGluGluLeuHisLysAsp 680
 QY 2240 GATGATGAGATGGCTTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAGAAG 2299
 Db AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
 QY 2300 GTCTGTCAGAGACTTCATGTCAGCTCATCAACACCCCACTCTCAACACGATGATCAG 2359
 Db ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
 QY 2360 TTCTGTGAACAGTTTGGAAAAGGACCGCAAAACACGTCGGCAAGCGCAGACATACC 2419
 Db PheCysGluGlnValTrpLysArgAspArgLysGlnAArgGlnAArgProGlyHisThr 740
 QY 2420 CCAGGGAACAGTAACAATGGAAGCACTTACAGAAATAAAGAAAGGTAGAAACAGGAGG 2479
 Db ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
 QY 2480 ACCCAGCAATTGAGAGGCGCCAGGAGTGC 2512
 Db ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 2

SM3A_RAT
 ID SM3A_RAT STANDARD; PRT; 772 AA.
 AC Q63548.
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 3A precursor (Semaphorin III) (Sema III).
 GN SEMA3A.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Brain;

RX MEDLINE=97073089; PubMed=8915837;
 RT Giger R.J., Wolfer D.P., De Wit G.M.J., Verhaagen J.;
 RA "Anatomy of rat semaphorin III/collapsin-1 mRNA expression and
 relationship to developing nerve tracts during neuroembryogenesis.";
 RL J. Comp. Neurol. 375:378-392(1996).
 CC -!- FUNCTION: MAY BE INVOLVED IN GUIDING GROWING AXONS TOWARDS THEIR
 CC TARGETS BY FORMING A MOLECULAR BOUNDARY THAT INSTRUCTS AXONS TO
 CC ENGAGE IN THE FORMATION OF SPECIFIC NERVE TRACTS. BINDS TO
 CC NEUROPILIN.
 CC -!- SUBCELLULAR LOCATION: Secreted (by similarity).
 CC -!- DEVELOPMENTAL STAGE: AT E-11, EXPRESSION WAS RESTRICTED TO THE
 CC OLFACTORY PIT, THE BASAL AND ROSTRAL SURFACE OF THE TELENCEPHALIC
 CC VESICLE, THE EYE ANLAGE, THE EPITHELIUM OF RATHKE'S POUCH, AND
 CC SOMITES. AT LATER DEVELOPMENTAL STAGES, IT WAS WIDELY DISTRIBUTED
 CC IN NEURONAL AS WELL AS IN MESENCHYMAL AND EPITHELIAL STRUCTURES
 CC OUTSIDE THE NERVOUS SYSTEM. AFTER BIRTH, MESENCHYMAL LEVELS
 CC DECREASED RAPIDLY AND EXPRESSION BECAME RESTRICTED TO SPECIFIC
 CC SETS OF NEURONS IN THE CNS. IN THE MATURE CNS, IT IS DETECTABLE IN
 CC MITRAL CELLS, NEURONS OF THE ACCESSORY BULB AND CEREBRAL CORTEX,
 CC CEREBELLAR PURKINJE CELLS, AS WELL AS A SUBSET OF CRANIAL AND
 CC SPINAL MOTONEURONS.
 CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
 CC THIRD OF THE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X95286; CAA64607.1; -.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 772 SEMAPHORIN 3A.
 FT DOMAIN 240 538 SEMA.
 FT DOMAIN 577 665 IG-LIKE C2-TYPE.
 FT DOMAIN 728 770 ARG/LYS-RICH (BASIC).
 FT DISULFID 650 723 BY SIMILARITY.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 591 591 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 772 AA; 88808 MW; 240907812PFPF2D2 CRC64;

Alignment Scores:

Pred. No.: 9.28e-316 Length: 772
 Score: 4057.50 Matches: 740
 Percent Similarity: 98.58% Conservative: 21
 Best Local Similarity: 95.85% Mismatches: 10
 Query Match: 82.87% Indels: 1
 DB: 1 Gaps: 1

US-09-774-490-1 (1-2709) x SM3A_RAT (1-772)

Qy 200 ATGGCGCTGGTTAACTAGGATGTCTCTTTCTGGGAGATATTACTACGCAAGACCA 259
 Db 1 MetGlyTrpPheThrGlyIleAlaCysLeuPheTrpGlyIleLeuLeuThrAlaArgAla 20


```

Db 741 SerGlnGlySerSerAsnLysTrpLysHisMetGlnGluSerLysLysGlyArgAsnArg 760
QY 2477 AGGACCCAGCAATTTGAGAGGCGCACCCAGGAGCTGTC 2512
Db 761 ArgThrHisGluPheGluArgAlaProArgSerVal 772

RESULT 3
SM3A_MOUSE
ID SM3A_MOUSE STANDARD; PRT; 772 AA.
AC O08665; Q62180; PRT; 772 AA.
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3A precursor (Semaphorin III) (Sema III) (Semaphorin D)
DE (Sema D).
GN SEMA3A OR SEMAD OR SEMD.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=NMRI; TISSUE=Embryo;
RX MEDLINE=95267431; PubMed=7748561;
RA Puschel A.W., Adams R.H., Betz H.;
RT "Murine semaphorin D/collapsin is a member of a diverse gene family
RT and creates domains inhibitory for axonal extension.";
RL Neuron 14:941-948(1995).
RN [2]
RC SEQUENCE FROM N.A.
RP MEDLINE=97470885; PubMed=9331345;
RX Taniguchi M., Yuasa S., Fujisawa H., Naruse I., Saga S., Mishina M.,
RA Yagi T.;
RT "Disruption of semaphorin III/D gene causes severe abnormality in
RT peripheral nerve projection.";
RL Neuron 19:519-530(1997).
RN [3]
RC SEQUENCE FROM N.A.
RP Kimura T., Fishman M.C.;
RT "cDNA sequence of mouse collapsin/semaphorin III.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RC SEQUENCE OF 107-772 FROM N.A.
RX TISSUE=Fetal brain;
RA Messersmith E.K., Leonardo E.D., Shatz C.J., Tessier-Lavigne M.,
RA Goodman C.S., Kolodkin A.L.;
RT "Semaphorin III can function as a selective chemorepellent to pattern
RT sensory projections in the spinal cord.";
RL Neuron 14:949-959(1995).
CC -!- FUNCTION: PLAY A ROLE IN GROWTH CONES GUIDANCE. MAY FUNCTION TO
CC NORMALLY TERMINATE DORSALLY.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- DEVELOPMENTAL STAGE: EXPRESSED EARLY IN EMBRYONIC DEVELOPMENT
CC (E11) IN DISTINCT REGIONS OF THE NEUROECTODERM AND MESODERM.
CC EXPRESSION BECAME MORE EXTENSIVE AT LATER STAGES.
CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
CC THIRD OF THE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and/or commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X85993; CAA59985.1; -

```

```

DR EMBL; D85028; BAA19773.1; -.
DR EMBL; L41541; AAL77611.1; -.
DR EMBL; L40484; AAA73934.1; -.
DR PIR; I48747; I48747.
DR PIR; I58169; I58169.
DR MGD; MGI:107558; Sema3a.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR001627; Sema.
DR Pfam; PF00047; IG; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS00835; IG-LIKE; 1.
KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 20
FT CHAIN 21 772
FT DOMAIN 240 538
FT DOMAIN 579 665
FT DOMAIN 728 770
FT DISULFID 650 723
FT CARBOHYD 53 53
FT CARBOHYD 125 125
FT CARBOHYD 591 591
FT CONFLICT 193 193
FT CONFLICT 207 207
FT CONFLICT 253 253
FT CONFLICT 352 352
FT CONFLICT 403 403
FT CONFLICT 571 572
FT CONFLICT 616 620
FT CONFLICT 623 623
SQ SEQUENCE 772 AA; 88799 MW; E89A08528B10AEC3 CRC64;

Alignment Scores:
Pred. No.: 6,5e-314 Length: 772
Score: 4035.50 Matches: 737
Percent Similarity: 98.32% Conservative: 22
Best Local Similarity: 95.47% Mismatches: 12
Query Match: 82.42% Indels: 1
DB: 1 Gaps: 1

US-09-774-490-1 (1-2709) x SM3A_MOUSE (1-772)
QY 200 ATGGCGTGGTAACTAGGATTCGTCTTTCTGGGAGTATTACTTACAGCAAGACA 259
Db 1 MetGlyTrpPheThrGlyIleAlaCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGATGGGAAGACAATGTCGAAGGCTGAATATCTCTACAAAGAATGTTG 319
Db 21 AsnTrpAlaAsnGlyLysAsnValProArgLeuLysLeuSerTrpLysGluMetLeu 40
QY 320 GAATCCCAATGATGATCACTTCAATGGCTGGCCCAACAGCTCCAGTTATCATACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTrpHisThrPhe 60
QY 380 CTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAGGATCATATTTTCATTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyValGlyAlaLysAspHisPheSerPhe 80
QY 440 GACTGTTTATCAAGGATTTTCAAAAGATTTGTGGCCAGTATCTTACACCAAGA 499
Db 81 AsnLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTrpArgArg 100
QY 500 GATCAATGCAAGTGGGCTGGAAAGACATCTGTAAGAAGATGTGCTAATTTTCATCAAGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATATACAGACTCACTTGTAGCCCTGTGGAAACGGGGGCTTTTTCATCCAATT 619

```

Db 121 LeuGluAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TCACCTACATTGAATTCGACATCATCTGAGCAATATTTTAAAGTCGAGAACTCA 679
Db 141 CysThrTyrIleGluValGlyHisProGluAspAsnIlePheLeuLeuGlnAspSer 160
QY 680 CATTTCGAAACGGCGTGGGAGAGTCATATGACCTTAAGCTCTGACAGCATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGGAGAAATTAATCTGGAACCTGAGCTGATTTTATGGGCGAGACTTTC 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTCCGAACTTCCTGGGCAACCAACCAATCAGACAGACAGATGATTCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCATGATCCAAAGTTCAATTAGTCCCACTCATCTCAGAGATGACAACTCCTGAAGT 919
Db 221 LeuAsnAspProArgPheIleSerAlaHisLeuIleProGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATCTTTCTTCCTGCGTGAATGCAATAGATGGAGAACACTCTCGAAAAGCT 979
Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTACGCTAGATAGGTAGATGCAAGATGACTTTGGAGGCAACAGATCTGCTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGACACATCTCTCAAGCTGCTGATTTGCTAGTCCAGGTCCAAATGCG 1099
Db 281 AsnLysTrpThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGATCAACTGCGAGATGATTCCTAATGAACTTTAAAGATCCTAAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnSerLysAspProLys 320
QY 1160 AATCCAGTTGTATAGGAGTGTTCAGACTTCAGTAACTTTCAAGGATCAGCCGTG 1219
Db 321 AsnProIleValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGATGATGAGAGGTGTTCTTGGTCCATATGCCACACAGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTATCCAGCGCAGGAAT 1339
Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCAGCAAAACATTTGGTGGTTTGATCTTCAAGAGGACTTCCTGATGATGTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACTTTGCAAGAGTCAATCCAGCATGATCAATCCAGTGTTCCTATGACATGCCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProIleAsnAsnArgPro 420
QY 1460 ATATGATCAAAACGGATGTAATATCAATTTACAAATTTGCTAGACCGAGTGGAT 1519
Db 421 IleMetIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAGATGAGACGATGATGTATGTTTATTCGGAACAGATGTTGGGACCTTCTTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 CTAGTTTCAATTCCTTAAGAGACTTGGTATGATTTAGAGAGTTCTGCTGGAAGATG 1639
Db 461 ValValSerValProLysGluThrTrpHisAspLeuLeuGluValLeuLeuGluMet 480
QY 1640 ACAGTTCCTGGGAACCGACTGCTATTTCAGCAATGGAGCTTTCCTAAGCAGCAACAA 1699

Db 481 ThrValPheArgGluProThrThrIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
QY 1700 CTATATATTGTTTCAACGGCTGGGTTCCCGAGCTCCCTTTACACCGGTGATATTTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGCGTGTGCTGAGTGTGCTGCCCGACACCTTACTGTCTGCTTGGATGGTCT 1819
Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCACTGCCAAGAGACGCAACAGACCAAGATATAAGAAAT 1879
Db 541 SerCysSerArgTyrPheProThrAlaLysArgThrArgArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACTCTACTGTTTCAGCTTA---CACCATGATATCACCATGGCCACAGC 1936
Db 561 GlyAspProLeuThrHisCysSerAspLeuGlnHisAspAsnHisGlyProSer 580
QY 1937 CCTCAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGCACATTTTTCGAATGCAGTCCG 1996
Db 581 LeuGluGluArgIleIleTyrGlyValGluAsnSerThrPheLeuGluCysSerPro 600
QY 1997 AAGTCGAGAGAGCGCTGCTTATTGGCAATTCAGAGCGCAATGAAGAGCGAAAGAA 2056
Db 601 LysSerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgAsnGluAspArgLysGlu 620
QY 2057 GAGATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCTCTCTCAGTAGTCTA 2116
Db 621 GluIleArgMetGlyAspHisIleArgThrGluGlnGlyLeuLeuLeuArgSerLeu 640
QY 2117 CAACAGAGAGATTCAGGCAATTTACTCTGCATCGGTGGAAACATGGGTTTCATACAACT 2176
Db 641 GlnLysLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheMetGlnThr 660
QY 2177 CTTCTTAAGTAAACCTCGAAGTCAATGACACAGAGCATTTGGAGAACTTCTTCATAAA 2236
Db 661 LeuLeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLys 680
QY 2237 GATCATGATGAGATGGCTCTTAAGACCAAGAAATGCTCAATAGCATGACACCTAGCCAG 2296
Db 681 AspAspAspGlyAspGlySerLysIleLysGluMetSerSerSerMetThrProSerGln 700
QY 2297 AAGTCTGTTACAGAGACTTCATGAGCTCATCAACACCCCAATCTCAACACAGATGGAT 2356
Db 701 LysValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAsp 720
QY 2357 GAGTCTCTGAAACAAGTTTGGAAAAGGACCGAAACAAACCTCGGCAAGCCAGGACAT 2416
Db 721 GluPheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHis 740
QY 2417 ACCCAGGGAACAGTAAACAAATGGAGCCTTACAGNAATTAAGAAAGTAGAACAGG 2476
Db 741 SerGlnGlySerSerAsnLysTrpLysHisMetGlnGluSerLysLysGlyArgAsnArg 760
QY 2477 AGGACCCACGAAATTTGAGAGGCGCCACCGAGGTGTC 2512
Db 761 ArgThrHisGluPheGluArgAlaProArgSerVal 772
RESULT 4
SM3A_CHICK
ID SM3A_CHICK STANDARD; PRT; 772 AA.
AC 090607;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3A precursor (Collapsin-1) (COLL-1).
GN SEMA3A OR COLL1.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 322-329; 362-372; 395-403 AND
 RC 666-680.
 RX TISSUE=Brain;
 RA MEDLINE=94006554; PubMed=8402908;
 RA Luo Y., Raible D., Raper J.A.;
 RT "Collapse: a protein in brain that induces the collapse and paralysis
 RT of neuronal growth cones.";
 RL Cell 75:217-227(1993).
 CC
 CC -!- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
 CC CONES. COULD SERVE AS A LIGAND THAT GUIDES SPECIFIC GROWTH CONES
 CC BY A MOTILITY-INHIBITING MECHANISM. BINDS TO NEUROPILIN.
 CC -!- SUBCELLULAR LOCATION: Secreted (by similarity).
 CC -!- TISSUE SPECIFICITY: EXPRESSED AT RELATIVELY HIGH LEVELS IN BRAIN
 CC AND MUSCLE, MODERATE LEVELS IN LUNG, BURSA, AND HEART AND
 CC VIRTUALLY ABSENT IN LIVER. COLLAPSEIN-1, -2, -3, AND -5 BIND TO
 CC OVERLAPPING BUT DISTINCT AXON TRACTS.
 CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
 CC THIRD OF THE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; U02528; AAC59638.1; -;
 CC PIR; A49069; A49069.
 CC InterPro; IPR007110; Ig-like.
 CC InterPro; IPR003599; Ig.
 CC InterPro; IPR003006; Ig_MHC.
 CC InterPro; IPR003659; Plexin-like.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF00047; Ig; 1.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00409; Ig; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC PROSITE; PS00835; IG_LIKE; 1.
 CC Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 CC Developmental protein; Glycoprotein.
 CC
 CC SIGNAL 1 22 POTENTIAL
 CC FT CHAIN 23 772 SEMAPHORIN 3A.
 CC FT DOMAIN 240 538 SEMA.
 CC FT DOMAIN 576 665 IG-LIKE C2-TYPE.
 CC FT DOMAIN 728 770 ARG/LYS-RICH (BASIC).
 CC FT DISULFID 650 723 BY SIMILARITY.
 CC FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SQ SEQUENCE 772 AA; 88867 MW; E91E09DE0CC940AC CRC64;
 CC
 CC Alignment Scores:
 CC Pred. No.: 6 91e-294 Length: 772
 CC Score: 3784.50 Matches: 683
 CC Percent Similarity: 95.34% Conservative: 53
 CC Best Local Similarity: 88.47% Mismatches: 35
 CC Query Match: 77.30% Indels: 1
 CC DB: 1 Gaps: 1
 CC
 CC US-09-774-490-1 (1-2709) x SM3A_CHECK (1-772)
 CC
 CC QY 200 ATGGGCTGTTAACTAGGATTGTCGTTCTTTCTGGGAGTATTACTTACACAGACGA 259
 CC
 CC Db 1 MetGlyTrpLeuArgGlyIleAlaLeuLeuSerLeuGlyValLeuAlaGlyArgVal 20
 CC
 CC QY 260 AACTATCAGAATGGGAGACAACTGTGCCAAGCTGAATATCTCTACAAAGATGTTG 319
 CC
 CC Db 21 AsnCysGlnHisValIysAsnAsnValProArgLeuLeuLysSerTyrIysGluMetLeu 40

QY 320 GAATCCAAACAATGTGATCACTTTCAATGGCTTGCCCAACAGCTCCAGTTATCATACCTTC 379
 Db 41 GluSerAsnAsnIleValAsnPheAsnGlyLeuAlaAsnSerSerSerfYHisThrPhe 60
 QY 380 CTTTGGATGAGGAACGAGTAGGCTGTATGTTGGAGCAAAAGGATCATCATATTTTCATTC 439
 Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
 QY 440 GACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGGCCAGTATCTTACACCAAGAAGA 499
 Db 81 AsnLeuValAsnIleLysGluTyrGlnLysIleValTrpProValSerHisSerArgArg 100
 QY 500 GATGAATCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTGCTAAATTTTCATCAAGSTA 559
 Db 101 AspGluCysLysTyrAlaGlyLysAspIleLeuArgGluCysAlaAsnPheIleLysVal 120
 QY 560 CTTAAGGATATATCAAGACTCATTGTGACGCTGTGGAACGGGGCTTTTCATCAATTT 619
 Db 121 LeuLysThrTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProMet 140
 QY 620 TGCACCTACATTGAAATTGGACATCATCTGAGGACAATATTTTAAAGCTGGAGAACTCA 679
 Db 141 CysThrTyrIleGluValGlySerHisProGluAspAsnIlePheArgMetGluAspSer 160
 QY 680 CATTTTGAAAACGGCGTGGGAAGAGTCCATATGACCCCTAAAGCTCTGACAGCATCCCTT 739
 Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
 QY 740 TTAATAGATGAGAAATTTATCTCTGGAACCTGAGTATTTTATGGGGGAGACTTGTCT 799
 Db 181 LeuValAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
 QY 800 ATCTTCCGAACTCTTTGGGCACACCACCAATCAGGACAGACAGCATGATTTCCAGGTGG 859
 Db 201 IlePheArgThrLeuGlyHisHisProLeuArgThrGluGlnHisAspSerArgTrp 220
 QY 860 CTCAATGATCCAAAGTTCATTAGTGGCCACCTCATCTCAGAGAGTGACAAATCTGGAAGAT 919
 Db 221 LeuAsnAspProArgPheIleSerAlaHisLeuIleProGluSerAspAsnProGluAsp 240
 QY 920 GACAAAGTATATCTTTTCTCCGTAATAATGCAATAGATGAGAGACACACTCGGAAAAGCT 979
 Db 241 AspLysIleTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisThrGlyLysAla 260
 QY 980 ACTCAGCTAGATAGGTGACATATGCAAGATGACTTTGGAGGCGACAGAGTCTGTGTG 1039
 Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
 QY 1040 AATAAATGGCAACAATTTCTCAAGCTCGTCTGATTGCTCAGTCCAGGTCCCAATGCG 1099
 Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
 QY 1100 ATTGACACTCAATTTTGATGAACCTGAGGATGTATTCCTTAATGAATTTAAAGATCCTAAA 1159
 Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnSerLysAspProLys 320
 QY 1160 AATCCAGTTGTATATGGAGTCTTTACGACTTCCAGTACATTTTCAAGGGATCACCGCTG 1219
 Db 321 AsnProIleValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
 QY 1220 TGTATGTATAGCATGATGTGAGAGGGTGTTCCTTGTGCTCATATGCCACACAGGAT 1279
 Db 341 CysMetTyrSerMetThrAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
 QY 1280 GGACCCAACTATCAATGGGTGCCCTTATCAAGAGAGAGTCCCTATTCACGGCCAGGAAC 1339
 Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
 QY 1340 TGTCCACGACAAACATTTGGTGGTTTACACTCTCAAGAGGACCTTCCTGATGATGTTATA 1399
 Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspGluValIle 400

1400 ACCTTTGCAAGAGTCATCCAGCCTGTACATCAATCCAGTGTTCCTTATGAAACAATCCGCCA 1459
Db |||||||
401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProIleAsnSerArgPro 420
Qy |||||||
1460 ATAGTGATCAAAACCGAGTAAATATCAATTTACACAAATGTCTGAGACCGAGTGGAT 1519
Db |||||||
421 IleMetIleLysThrAspValAspTyrGlnPheThrGlnIleValValAspArgValAsp 440
Qy |||||||
1520 GCAGAAGATGGACAGTATGATTTATCGGACAGATGTTGGACCGCTTCTTAAA 1579
Db |||||||
441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspIleGlyThrValLeuLys 460
Qy |||||||
1580 GTAGTTTCAATTCCTAAGCAGACTGTGTATGATTTAGAGAGGTCTGCTGGAAGAAATG 1639
Db |||||||
461 ValValSerIleProLysGlnThrPheHisGluLeuGluValLeuLeuGluGluMet 480
Qy |||||||
1640 ACAGTTTTCGGGAACCGACTGCTATTTCCAGCAATGGAGCTTCCCAATGAGCAGCAACAA 1699
Db |||||||
481 ThrValPheArgGluProThrValIleSerAlaMetLysIleSerThrLysGlnGlnGln 500
Qy |||||||
1700 CTATATATTTGGTTCACGCTGGGTGCCAGCTCCCTTTACACGGGTGCTATTTAC 1759
Db |||||||
501 LeuTyrIleGlySerAlaThrGlyValSerGlnLeuProLeuHisArgCysAspValTyr 520
Qy |||||||
1760 GCGAAGCGTGTGCTGAGTGTTCCTCCGCGCAGACCTTACTGTGCTGGATGTTCT 1819
Db |||||||
521 GlyLysAlaCysAlaGluCysLysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
Qy |||||||
1820 GCATGTTCTCGTATTTTCCCACTGCAAGAGACGACGACGACGACGACGACGACGACG 1879
Db |||||||
541 SerCysSerArgTyrPheProThrAlaLysArgThrArgGlnAspIleArgAsn 560
Qy |||||||
1880 GGAGACCCACTGACTCAGTGTTCAGACTTA---CACCATGATTAATCACCATGCCACAGC 1936
Db |||||||
561 GlyAspProLeuThrHisCysSerAspLeuGlnHisHisAspAsnProSerGlyGlnThr 580
Qy |||||||
1937 CCTGAAGAGAGATCATCTATGCTGTAGAGATAGTAGACACATTTTGGAAATCGAGTCG 1996
Db |||||||
581 LeuGluGluLysIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerPro 600
Qy |||||||
1997 AAGTCGCAGAGCGCTGCTATTGGCAATTCAGAGCGCAAAATCAAGAGCGCAAAAGAA 2056
Db |||||||
601 LysSerGlnArgAlaIleValTyrTrpGlnPheGlnLysGlnAsnAspHisLysVal 620
Qy |||||||
2057 GAGATCAGAGTGATGATCATATCATCAGCAGATCAAGCGCTTCTGCTAGTGTCTA 2116
Db |||||||
621 GluIleLysValAspAspArgMetIleArgThrGluGlnGlyLeuLeuLeuArgSerLeu 640
Qy |||||||
2117 CAACAGAGCATTCAGGCAATACCTCTGCCATGCGGTGGAACATGGGTTCATACAACT 2176
Db |||||||
641 GlnArgAspSerGlyIleTyrPheCysHisAlaValGluHisGlyPheIleGlnThr 660
Qy |||||||
2177 CTTCTTAAGTTAACCTCGAAGTCATTTGACACAGACATTTTGGAAAGAACTTCTTCAAAA 2236
Db |||||||
661 LeuLeuLysValThrLeuGluValIleAspThrAspHisLeuGluGluLeuHisLys 680
Qy |||||||
2237 GATGATGATGAGATGGCTCTAAGACCAAGAAATGTCCAATAGATGACCTAGCCAG 2296
Db |||||||
681 GluGluAspAlaSerLysThrLysAspAlaThrAsnSerMetThrProSerGln 700
Qy |||||||
2297 AAGTCTGCTGACAGACTTCATGACGCTCATCAACCCCACTCAACAGCATGGAT 2356
Db |||||||
701 LysIleTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAsp 720
Qy |||||||
2357 GAGTCTGTGAACAAGTTTGGAAAGGACCGGAAACAACAGCTCGGCAAGCCGACGACAT 2416
Db |||||||
721 GluPheCysGluGlnValTrpLysArgAspArgLysGlnArgArgGlnArgProAlaAsn 740
Qy |||||||
2417 ACCCGAGGACAGTAAACAAATGGAAGCACTTACAGAAATAAGAAAGTGAACAGCAGG 2476
Db |||||||
741 AlaGlnValAsnThrAsnLysTrpLysHisLeuGlnGlnAsnLysLysGlyArgAsnArg 760
Qy |||||||
2477 AGGACCCAGCAATTTGAGAGGCGCACCCAGGAGTGTC 2512

Db 761 ArgThrHisGluPheGluArgAlaProArgSerVal 772

RESULT 5

SZ1B BRARE
ID SZ1B BRARE STANDARD; PRT; 778 AA.
AC Q9W686;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin z1b precursor (Semaphorin 1B) (Sema-z1b).
GN SEMAZ1B OR SEMAZ3AB.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99425174; PubMed=10495275;
RA Roos M., Schachner M., Bernhardt R.R.;
RT "Zebrafish semaphorin z1b inhibits growing motor axons in vivo.";
RL Mech. Dev. 87:103-117(1999).
CC -!- FUNCTION: MIGHT NORMALLY INFLUENCE THE MIDSEGMENTAL PATHWAY CHOICE
OF THE VENTRALLY EXTENDING MOTOR AXONS BY CONTRIBUTING TO A
REPULSIVE DOMAIN IN THE POSTERIOR SOMITE.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- TISSUE SPECIFICITY: EXPRESSED IN RHOMBOMERES THREE AND FIVE, AND
IN THE POSTERIOR HALF OF NEWLY FORMED SOMITES WHICH IS AVOIDED BY
VENTRALLY EXTENDING MOTOR AXONS.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AF083382; AAD28103.1; -
ZFIN; ZDB-GENE-991209-6; sema3ab.
InterPro; IPR007110; Ig-like.
InterPro; IPR003599; Ig.
InterPro; IPR003006; Ig_MHC.
InterPro; IPR003659; Plexin-like.
InterPro; IPR002165; Plexin_repeat.
InterPro; IPR001627; Sema.
Pfam; PF00047; Ig_1.
Pfam; PF01437; PSI; 1.
Pfam; PF01403; Sema; 1.
SMART; SM00409; IG; 1.
SMART; SM00423; PSI; 1.
SMART; SM00630; Sema; 1.
PROSITE; PS50835; IG_LIKE; 1.
Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 17
FT CHAIN 18 778
FT SEMAPHORIN Z1B.
FT DOMAIN 241 539
FT SEMA.
FT DOMAIN 579 668
FT IG-LIKE C2-TYPE.
FT DOMAIN 721 776
FT ARG/LYS-RICH (BASIC).
FT DISULFID 652 716
FT CARBOHYD 54 54
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 127 127
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 593 593
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 778 AA; 88904 MW; 4D36F4323AE21895 CRC64;

Alignment Scores:

Pred. No.: 7,49e-248 Length: 778
Score: 3207.50 Matches: 591

Percent Similarity:	84.90%	Conservative:	78
Best Local Similarity:	75.00%	Mismatches:	92
Query Match:	65.51%	Indels:	27
DB:	1	Gaps:	8

US-09-774-490-1 (1-2709) x SZ1B_BRARE (1-778)

Qy	200	ATGGGCTGTTAACTAGGATTGTCGTCTTTCTTGGGAGTATTACTTACACGAAGACA	259
		:	:
Db	1	MetAspTyrLeuTrpIleValLeuLeuIleTrpThrLeuIleAalaProGluArgGly	20
		:	:
Qy	260	AACATAT---CAGAATGGCAAGAACAAATGTCACAGCTGAAATATCTCTACAAAGAAATG	316
		:	:
Db	21	ThrVallalaGlnA-gSerLysSerAsnValProArgLeuLysProSerTyrLysGluMet	40
		:	:
Qy	317	TTGGAATCCAAACATATGTCATCTTCAATGGCTTGGCCACACAGCTCCAGCTTATCATACC	376
		:	:
Db	41	LeuGluSerAsnAsnLeuLeuThrPheAsnGlyLeuAlaAsnSerSerAlaTyrHisThr	60
		:	:
Qy	377	TTCCCTTTTGGATGAGGAACCGAGTAGGCTGTATGTTGGAGCAAAAGGATCACATATTTTCA	436
		:	:
Db	61	PheLeuLeuAspGluGluArgGlyArgLeuPheValGlyAlaLysAspHisValLeuSer	80
		:	:
Qy	437	TTCCACCTGGTTAATATCAAG--GATTTTCAAAGATTGTGTGGCCAGTATCTTACACC	493
		:	:
Db	81	PheAsnLeuValAspIleAsnMetAspGlnGlnLeuIleSerTrpProSerSerProSer	100
		:	:
Qy	494	AGAAGAGATGAATGCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTCTAAATTTTCATC	553
		:	:
Db	101	ArgargaspGluCysLysTrpAlaGlyLysaspValGlnLysGluCysAlaAsnPhelle	120
		:	:
Qy	554	AAGTACTTAAGCATATAATCAGACTCACTTGTCACGCTCTGGAACGGGGCTTTTTCAT	613
		:	:
Db	121	LysValLeuGlnProPheAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHis	140
		:	:
Qy	614	CCAAITTTGCRCCTACATGAATTTGGACATCATCTCTGAGGACAATAATTTTAAAGCTGAG	673
		:	:
Db	141	ProValCysAlaHisValGluValGlyLysargSerGluAspAsnThrPheArgLeu---	159
		:	:
Qy	674	AACTCACATTTTGAACCGCGCTGGGAAGAGTCCATATGACCTCAAGCTGCTGACAGCA	733
		:	:
Db	160	GlySerSerPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuGlnThrAla	179
		:	:
Qy	734	TCCCTTTTAATAGATGGAGAATATACTCTGAACTGCAGCTGATTTTATGGGCGAGAC	793
		:	:
Db	180	SerMetLeuIleAspGlyGluLeuTyrAlaGlyThrSerAlaAspPheMetGlyArgAsp	199
		:	:
Qy	794	TTTGCTATCTCCGAATCTTGGGCACACCACCCCAATCAGACAGACAGCATGATTCC	853
		:	:
Db	200	PheAlaIlePheArgThrLeuGlyLysHisHisProIleArgThrGluGlnHisAspSer	219
		:	:
Qy	854	AGTGGCTCAATGATCCAAAGTTCATATGATGCCACCTCATCTCAGAGAGTGACAATCCT	913
		:	:
Db	220	ArgTrpleuAsnAspProArgPheValHisLeuIleProGluSerAspAsnAla	239
		:	:
Qy	914	GAAGATGACAAAGTATATCTTTTCTCCGTGAAAATGCAATAGATGGAGAACACTCTGGA	973
		:	:
Db	240	GluAspAspLysIleTyrLeuPhePheArgGluAsnAlaIleAspGlyGluGlnIleSer	259
		:	:
Qy	974	AAAGCTACTCACGCTAGAAATAGTTCAGATATGCAAGAAATGACTTTTGGAGGGCACAGAGT	1033
		:	:
Db	260	LysAlaThrHisAlaArgIleGlyGlnLeuCysLysAsnAspPheGlyGlyHisArgSer	279
		:	:
Qy	1034	CTGGTGAATAAATCGACAACATCTCTCAAAGCTCGTGTGATTTCTCAGTCCAGGTCCTCA	1093
		:	:
Db	280	LeuValAsnLysTrpThrPheLeuLysAlaAlaArgLeuValCysSerValProGlyLeu	299
		:	:
Qy	1094	AATGGCATTCACACTCATTTTTGATGAACATCGAGGATGATTTCTCTAAATGAACTTTAAAGAT	1153
		:	:
Db	300	AsnGlyIleAspThrHisPheAspGluLeuGlnAspValPheLeuMetSerSerLysAsp	319
		:	:
Qy	1154	CCTAAAATCCAGTTGTATATGGAGTGTTTACGACTTCCAGTAACTTTTCAAGGGATCA	1213
		:	:

Db 498 GlnIlePheLysHisSerSerIleIleLeuLeuMetGluLeuSerLeuLysGlnGln 517
 QY 1700 TATATATTGGTTCAACGGCTGGGTTGCGCCAGCTCCCTTTACACCGGTGTGATATTAC 1759
 Db 518 LeuTyrlleGlySerArgAspGlyLeuValGlnLeuSerLeuHisArgCysAspThrTyr 537
 QY 1760 GGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1819
 Db 538 GlyLysAlaCysAlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyAsn 557
 QY 1820 GCATGTTCTCGCTATTTCCCACTCCAAAGAGAGCCACAGACCAAGATATAGAAAT 1879
 Db 558 AlaCysSerArgTyrAlaProThrSerLysArgAlaArgGlnAspValLysTyr 577
 QY 1880 GGAGACCACTGACTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1939
 Db 578 GlyAspProIleThrGlnCysTrpAspIleGluAspSerIleSerHis--GluThrAla 596
 QY 1940 GAAGAGAGATCATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1999
 Db 597 AspGluLysValIlePheGlyIleGluPheAsnSerThrPheLeuGluCysIleProLys 616
 QY 2000 TCGCAGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2059
 Db 617 SerGlnGlnAlaThrIleLysTrpTyrIleGlnArgSerGlyAspGluHisArgGluGlu 636
 QY 2060 ATCAGAGGATGATCATATCATCATCAGCAGACATCAAGCGCTTCTGCTAGTGTCTACAA 2119
 Db 637 LeuLysProAspGluArgIleIleLysThrGluTyrGlyLeuLeuIleArgSerLeuGln 656
 QY 2120 CAGAGAGATTGAGCAATTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2179
 Db 657 LysLysAspSerGlyMetTyrTyrCysLysAlaGlnGluHisThrPheIleHisThrIle 676
 QY 2180 CTTAAGGTAACCTCGAAGTCTATGATCATCATCAGCAGACATCAAGCGCTTCTGCTAGTGTCT 2239
 Db 677 ValLysLeuThrLeuAsnValIleGluAsnGluGlnMetGluAsnThrGlnArgAlaGlu 696
 QY 2240 GATGATGAGATGCTCTAAGACCAACCAAGAAATGCTCAATGATGATGATGATGATGATGAT 2299
 Db 697 His-----GluGluGlyLysValLysAsp-----LeuLeuAlaGluSerArg 710
 QY 2300 GTCTGTTACAGAGACTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2359
 Db 711 LeuArgTyrLysAspTyrIleGlnIleLeuSerSerProAsnPhe---SerLeuAspGln 729
 QY 2360 TTCTGTGAACAGTTTGGAAAGGAGCCGCAACCAACCGTCCGCAAGCCAGGACATACC 2419
 Db 730 TyrCysGluGlnMetTrpHisArgGluLysArgGlnArgAsnLysGlyGly----- 747
 QY 2420 CCAGGGAACAGTAACAATGGAAGCATTCAAGAAATAGAAAGGTAGAAACAGGAGG 2479
 Db 748 -----ProLysTrpLysHisMetGlnGluMetLysLysLysLysLysLysLysLysLys 763
 QY 2480 ACCAC---GAATTGAGAGGCGCCAGGAGTGTC 2512
 Db 764 HisHisArgAspLeuAspGluLeuProArgAlaVal 775

RESULT 8

SM3D_CHICK
 ID_SM3D_CHICK STANDARD; PRT; 761 AA.
 AC Q90663;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Senaphorin 3D precursor (Collapsin-2) (COLL-2).
 GN SEN3D OR COLL2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]

US-09-774-490-1 (1-2709) x SM3D_CHICK (1-761)

QY 188 AGCGTCTGACGATGGCTGTTAACTAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 247
 Db 8 AshAlaCysSer-----LeuLeuSerLeuAlaMetLeuPheProValThrGly 24
 QY 248 ACAGCAAGACAACTATCATGATGGAAGCAACATGTCAGAGCTGCAATTTATCTCTAC 307
 Db 25 ThrSer-----LysGlnAsnIleProArgLeuLysLeuSerTyr 37
 QY 308 AAAGAATGTTGAATCCAAATGATGATCACTTTCAATGGCTTGGCCACAGCTCCAGT 367
 Db 38 LysAspLeuLeuSerAsnSerCysIleProPheLeuGlySerThrGluGlyLeuAsp 57
 QY 368 TATCATACCTTCCTTTTGGATGAGGAACGAGTAGGCTGTATGTTGGAGCAAGGATCAC 427

Alignment Scores:
 Pred. No.: 4, 59e-171 Length: 761
 Score: 2245.00 Matches: 414
 Percent Similarity: 72.72% Conservative: 151
 Best Local Similarity: 53.28% Mismatches: 182
 Query Match: 45.85% Indels: 30
 DB: 1 Gaps: 12

SEQUENCE FROM N.A.
 MEDLINE=95329269; PubMed=7605628;
 Luo Y., Shepherd I., Li J., Renzi M.J., Chang S., Raper J.A.;
 "A family of molecules related to collapsin in the embryonic chick nervous system";
 Neuron 14:1131-1140 (1995).
 CC - FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH CONES. COULD POTENTIALLY ACT AS REPULSIVE CUES TOWARD SPECIFIC NEURONAL POPULATIONS. BINDS TO NEUROPILIN.
 CC - SUBCELLULAR LOCATION: Secreted.
 CC - TISSUE SPECIFICITY: DEVELOPING SPINAL CORD AND DEVELOPING VISUAL SYSTEM. COLLAPSIN-1, -2, -3, AND -5 BIND TO OVERLAPPING BUT DISTINCT AXON TRACTS.
 CC - DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY THIRD OF THE PROTEIN.
 CC - SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC - SIMILARITY: Contains 1 Sema domain
 CC - SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 EMBL: U28240; AAA86896.1; --
 InterPro: IPR007110; Ig-like.
 InterPro: IPR003599; Ig.
 InterPro: IPR003006; Ig_MHC.
 InterPro: IPR003659; Plexin-like.
 InterPro: IPR001627; Sema.
 Pfam: PF01403; Sema; 1.
 SMART: SM00409; IG; 1.
 SMART: SM00423; PSI; 1.
 SMART: SM00630; Sema; 1.
 PROSITE: PS50835; IG_Like; 1.
 Signal: Immunoglobulin domain; Multigene family; Neurogenesis; Developmental protein; Glycoprotein.
 SIGNAL 1 24 POTENTIAL.
 CHAIN 25 761 SEMAPHORIN 3D.
 DOMAIN 245 543 SEMA.
 FT DOMAIN 552 670 IG-LIKE C2-TYPE.
 FT DOMAIN 727 757 ARG/LYS-RICH (BASIC).
 FT DISULFID 653 719 BY SIMILARITY.
 FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 595 595 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 761 AA; 87300 MW; 3E09AE3DBA53F46B CRC64;

AC O9W6G6;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DE 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Semaphorin Z2 precursor (Semaphorin 2) (Sema-Z2).
 GN SEMA22 OR SEMA2.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99112778; PubMed=9915572;
 RA Halloran M.C., Severance S.M., Yee C.S., Gemza D.L., Raper J.A.,
 RA Kuwada J.Y.;
 RT "Analysis of a Zebrafish semaphorin reveals potential functions in
 RT vivo";
 RL Dev. Dyn. 214:13-25(1999).
 CC -!- FUNCTION: MAY PLAY A ROLE IN THE GUIDANCE OF SEVERAL AXON
 CC PATHWAYS.
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN A DYNAMIC AND RESTRICTED PATTERN
 CC DURING THE PERIOD OF AXON OUTGROWTH.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF124485; RAD21310.1; --
 CC ZFIN; ZDB-GENE-990715-2; sema2.
 CC InterPro; IPR003599; Ig.
 CC InterPro; IPR003659; Plexin-like.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00409; IG; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 CC Developmental protein; Glycoprotein.
 CC SIGNAL 1 41 POTENTIAL.
 CC CHAIN 42 764 SEMAPHORIN Z2.
 CC DOMAIN 261 559 SEMA.
 CC DOMAIN 661 740 IG-LIKE C2-TYPE.
 CC DOMAIN 741 762 ARG/LYS-RICH (BASIC).
 CC DISULFID 668 733 BY SIMILARITY.
 CC CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 490 490 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC CARBOHYD 610 610 N-LINKED (GLCNAC. . .) (POTENTIAL).
 CC SEQUENCE 764 AA; 87859 MW; A3BD95C2C479D7AE CRC64;
 SQ
 Alignment Scores:
 Pred. No.: 1-72e-161 Length: 764
 Score: 2125.00 Matches: 390
 Percent Similarity: 71.37% Conservative: 156
 Best Local Similarity: 50.98% Mismatches: 191
 Query Match: 43.40% Indels: 28
 DB: 1 Gaps: 11
 US-09-774-490-1 (1-2709) x SMZ2_BRARE (1-764)
 QY 215 AGGATTGCTGCTTTCTGGGAGTATTATCTTACGACGAGCAAACTATCAGAATGGG 274
 Db 20 ArgPheSerCysAlaTrpTrpSerThrSerValMetLeuPhePheSerLeuProGluGly 39
 QY 275 -----AGAACAAATGTGCCAGGCTGAAATTTATCTCTACAAAGAAATGTTGAATCC 325

Db 40 AsnCysMetLysGluSerLeuProArgValLysLeuGlyTyrLysAspLeuLeuHisSer 59
 QY 326 AACAAATGTGATCACCCTTTCAATGGCTTGGCCACAGCTCCAGTTATCATCACTCTCTTTG 385
 Db 60 ArgSerValProPheThrGlySerSerGluGlyGlnHisPheGlnThrValLeuLeu 79
 QY 386 GATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAGGATCATATTTTTCATTCGACCTG 445
 Db 80 AspGluGluArgSerArgLeuLeuGlyAlaLysAspHisValTyrLeuLeuAspPro 99
 QY 446 GTTAATATC---AAGGATTTTCAAGATTTGTGTCGCCAGTATCTTACACGAGAGAT 502
 Db 100 AspAsnIleAsnLysHisProLysLysLeuSerTrpProAlaSerArgAspArgValGlu 119
 QY 503 GAATCAAGTGGCTGGAAAGACATCTCTGAAAGATGTGCTAAATTTTCATCAAGTACTT 562
 Db 120 MetCysIleLeuAlaGlyLysAsnProLeuThrGluCysAlaAsnPheIleArgValLeu 139
 QY 563 AAGCATATAATCAGACTCCTGTGACGCTGTGGAACGGGGCTTTTCATCCAAATTTGC 622
 Db 140 HisSerTyrAsnArgThrHisValTyrAlaCysGlyThrGlyAlaPheHisProThrCys 159
 QY 623 ACCTACATTGAATTTGGACATCCTCTGAGACAATATTTTAACTGGAGAACTCATCAT 682
 Db 160 AlaPheLeuGluIleLysGlyHisLysGluAspArgTrpLeuLeuLeuHisSerAsnThr 179
 QY 683 TTTGAAACGGCTGGGAGAGAGTCCATATGACCCCTAAGCTGCTCAGACGATCCCTTTTA 742
 Db 180 MetGluSerGlyArgMetLysCysProPheAspProAsnGlnProPheAlaSerValLeu 199
 QY 743 ATAGATGAGAAATATATCTCTGGAACCTGAGTATTTTATGGGCGAGAGACTTTGCTATC 802
 Db 200 ThrAspGlnTyrLeuTyrAlaGlyThrAlaSerAspPheLeuGlyLysAspSerThrPhe 219
 QY 803 TTCGAACTCTTGGG-----CACCACACCCCAATCAGACGACGAGCATGATCC 853
 Db 220 ThrArgSerLeuGlyProProHisGlnGlnTyrIleArgThrAspIleSerGluAsp 239
 QY 854 AGTGGCTCAATGATCCAAAGTTCTATTAGTGCACCTCATCTCAGAGAGTGACAACTCT 913
 Db 240 TyrTrpIleAsnGluGlyLysPheIleSerAlaHisProIleSerAspThrTrpAsnPro 259
 QY 914 GAAGATGACAAATATATCTTTCTCGTGAAATGCAATAGATGAGAGACACTCTCGA 973
 Db 260 AspAspAspLysIleTyrPhePheArgGluAlaSerArgAspGlySerThrThrAsp 279
 QY 974 AAGCTACTCAGCTAGATAGGTGAGATATGCAAGATGATCTTGGAGGGCAGCAAGT 1033
 Db 280 LysSerValLeuSerArgValAlaArgIleCysArgAsnAspValGlyLeuArgSer 299
 QY 1034 CTGGTGAATAATGACAACTCTCTCAAGCTGCTGATTTGCTGATGCTAGTGCCAGGTCCA 1093
 Db 300 LeuThrAsnLysTrpThrPheLeuLysAlaArgLeuValCysSerIleProGlyPro 319
 QY 1094 AATGGCATTCACACTCATTGTCGAACTGAGGATGATTCCTTAATGAACCTTTAAAGAT 1153
 Db 320 AspGlyValAspThrHisPheAspGluLeuGlnAspIlePheLeuLeuProSerArgAsp 339
 QY 1154 CCTAAATTCAGTTGATATGAGAGTGTTCAGCTCCAGTCCAGTAACATTTTCAGGATCA 1213
 Db 340 GluLysAsnProMetValTyrGlyValPheThrThrSerSerIlePheLysGlySer 359
 QY 1214 GCCGTGTGTATGATAGCATGAGTGTGAGAGGGTTCCTTGGTCCATATGCCAC 1273
 Db 360 AlaValCysValTyrThrMetGluAspIleArgAlaAlaPheAsnGlyProTyrAlaHis 379
 QY 1274 AGGATGACCACTATCATGCTGCTTATCAAGAGAGAGTCCCTATCCACGGCA 1333
 Db 380 LysGluGlyProAspHisArgTrpValGluTyrGluGlyArgGlyIleProTyrProArgPro 399
 QY 1334 GGAATCTGTCCACAGAAACATTTGGT---GGTTTGACTCTACAAAGAGACTTCCTCAT 1390
 |||||

Db 400 GlyThrCysProSerArgThrTyrAspProHisIleLeuThrThrLysAspPheProAsp 419
QY 1391 GATGTTATACCTTTTCAGAGATCATCCAGCATATCAATCCAGTGTTCCTATGAC 1450
Db 420 GluValIleSerPheIleArgLeuHisProLeuMetTyrGlnSerValHisProMetThr 439
QY 1451 AATCGCCCAATAGTATCAAAAGCGATGTAATTTATCAATTTACAAATTTGCTGAGAC 1510
Db 440 GlyArgProIlePheThrArgIleAsnThrGluTyrArgLeuThrGlnIleIleValAsp 459
QY 1511 CGAGTGGATCGCAAGATCGACAGTATGATGTTATGTTATCGGACAGATGTTGGACC 1570
Db 460 ArgValAlaAlaGluAspGlyGlnTyrAlaValMetPheLeuGlyThrAspMetGlySer 479
QY 1571 GTTCTTAAAGTAGTTTCAATTCCTAAGGAGACTGTGTATGATTTAGAGAGTTCCTGCTG 1630
Db 480 ValLeuLysValValSerIleThrGlnGluAsnTrp---SerSerGluGluIleIleLeu 498
QY 1631 GAAGAATACACAGTTCCTCGGAACCGACTGCTATTTCAGCAATGAGCTTTCCTAAG 1690
Db 499 GluGluLeuGlnValPheLysAsnProSerProIleLeuAsnMetGluValSerSerLys 518
QY 1691 CAGCAACAATATATGTTTCAACGGCTGGGTTGCCAGCTCCCTTTACACCGGTGT 1750
Db 519 GlnGlnGlnLeuPheValGlyGlySerAspGlyLeuValGlnValSerLeuHisArgCys 538
QY 1751 GATATTACGGAAACGGTGTCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1810
Db 539 GlnIleTyrGlyGlnGlyCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrp 558
QY 1811 GATGGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1870
Db 559 AspGlyThrGlnCysSerArgTyrIleProAlaSerLysArgAlaArgGlnAsp 578
QY 1871 ATAAGAATGGAGACCATCATCTACTGTTACAGACTTACACCATGATATCACCATGCG 1930
Db 579 IleLysHisGlyAspProSerHisCysTrpAspThr-----GluAspValLeuGly 596
QY 1931 CACAGCCCTGAAGAGATCATCTATGTTGATAGAAATAGTAGACATTTTGGAAATGC 1990
Db 597 ArgAsnValGluGlnLysValLeuTyrGlyValGluSerAsnSerPheLeuGluCys 616
QY 1991 AGTCGCAAGTCGAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2050
Db 617 ValSerLysSerGlnGlnAlaLeuIleArgTyrPheValLeuLysProGlyValAspHis 636
QY 2051 AAAGAGATCATGATGATCATATCATCAGGACAGATCAAGGCTTCTGTACTGT 2110
Db 637 ArgGlnGluIleLysProAspGluArgValLeuIleThrAspArgGlyLeuLeuIleArg 656
QY 2111 AGTCTACACAGAGATTCAGGCATTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2170
Db 657 TrpLeuGlnArgGlyAspAlaGlySerTyrPheCysThrSerGlnGlnHisArgPheThr 676
QY 2171 CAACCTCTTCTTAAGTACCTCGAAGTCATTGACACAGAGATTTGGAAGAACTTCT 2230
Db 677 ArgThrLeuLeuHisValSerLeuHisIleLeuAspArgGlyClnIleAsn-----Ala 694
QY 2231 CATAAGATGATGATGATGATGCTTCAAGACAAAGAAATGCTCAATAGCATGACCT 2290
Db 695 HisGlnProAlaIleArgGluSerSerGluAsnProAlaValThrGlu-----Pro 711
QY 2291 AGCCAGAAGTCTGTGTACAGAGATTCATGCAGCTCATCAACACCCCAATCTCAACG 2350
Db 712 ArgGlnArg-----TyrLysAspTyrLeuArgMetLeuSerGlyPro---AlaArgSer 728
QY 2351 ATGGATGATGATGATGATGATGCTTCAAGACAAAGGACCGGAAACCAACGCGCAAGG 2410
Db 729 LeuAspGluTyrCysGluThrMetTrpHisArgGluLysGlnLysGlnLys----- 746
QY 2411 GGACATACCCCGGGAACAGTAAATGAGAGCATTAACAAGAAATGAAGAAAGTACA 2470
Db 747 -----GlyLysTrpLysHisValGlnGluLeuArgLysSerArg 759

QY 2471 AACAGGAGGCCAC 2485
Db 760 AsnArgArgHisHis 764
RESULT 10
SM3B HUMAN
ID SM3B HUMAN STANDARD; PRT; 749 AA.
AC Q13214; Q8TDV7; Q93018;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3B precursor (Semaphorin V) (Sema V) (Sema A(V)).
GN SEMA3B OR SEMA5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96210603; PubMed=8633026;
RA Sekido Y., Bader S., Latif F., Chen J.-Y., Duh F.-M., Wei M.-H.,
RA Albanesi J.P., Lee C.-C., Lerman M.I., Minna J.D.;
RT "Human semaphorins A(V) and IV reside in the 3p21.3 small cell lung
RT cancer deletion region and demonstrate distinct expression patterns.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:4120-4125(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX Dante M., Wamsley P.;
RA Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX Koyama N.;
RA "Semaphorin 3B (SEMA3B) cDNA";
RT Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP VARIANTS NSCLC CVS-348; HIS-397 AND ILE-415.
RX MEDLINE=20535986; PubMed=11085536;
RA Lerman M.I., Minna J.D.;
RT "The 630-Kb lung cancer homozygous deletion region on human chromosome
RT 3p21.3: identification and evaluation of the resident candidate tumor
RT suppressor genes";
RL Cancer Res. 60:6116-6133(2000).
CC -!- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
CC SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY). ACCUMULATES IN THE
CC ENDOPLASMIC RETICULUM.
CC -!- TISSUE SPECIFICITY: Expressed abundantly but differentially in a
CC variety of neural and nonneural tissues.
CC -!- DISEASE: Defects in SEMA3B are found in non-small cell lung cancer
CC (NSCLC) cell lines.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U28369; AAD09138.1; -;
CC EMBL; U73167; AAC02731.1; -;
CC EMBL; AB083186; BAB88870.1; -;
CC PIR; G01856; G01856.
CC Genew; HGNC:10724; SEMA3B.
CC MIM; 601281; -;
CC GO; GO:0005783; C:endoplasmic reticulum; TAS.
CC GO; GO:0007411; P:axon guidance; TAS.
CC GO; GO:0007267; P:cell-cell signaling; TAS.

DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF0047; Ig; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; Ig; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS50835; IG LIKE; 1.
 KW Signal; Immunoglobulin domain; Multigene family; Glycoprotein;
 KW Disease mutation.
 FT SIGNAL 1 24
 FT CHAIN 25 749
 FT DOMAIN 55 497
 FT DOMAIN 573 659
 FT DOMAIN 699 702
 FT DOMAIN 724 744
 FT DISULFID 644 710
 FT CARBOHYD 82 82
 FT CARBOHYD 124 124
 FT CARBOHYD 427 427
 FT VARIANT 348 348
 FT VARIANT 397 397
 FT VARIANT 415 415
 FT CONFLICT 29 36
 FT CONFLICT 308 308
 SQ SEQUENCE 749 AA; 83121 MW; 1F3B8F63F59444F3 CRC64;

Alignment Scores:
 Pred. No.: 6.2e-161
 Score: 2118.00
 Percent Similarity: 67.2%
 Best Local Similarity: 51.53%
 Query Match: 43.26%
 Indels: 72
 Gaps: 12

US-09-774-490-1 (1-2709) x SM3B_HUMAN (1-749)

QY 221 GTCGTCTTTCTGGGAGTATTACTACAGCAGAGCAAACTATCAGATGGGAAGAAC 280
 Db 13 LeuAlaLeuLeuTrpAlaValGlyLeuGlySerAlaAla-----Pro 26
 QY 281 AATGTCGAAGCTGAAATATCTTACAAAGAAATGTTGGATCCAAACATGTGATCACT 340
 Db 27 SerProProArgLeuArgLeuSerPheGlnGluLeuGlnAlaTrpHisGlyLeuGlnThr 46
 QY 341 TTCAATGCTTGGCCCAACAGCTCCAGTTATCATACCTTCTCTTGGATGAGGACGGAGT 400
 Db 47 PheSer---LeuGluArgThrCysCysTyrGlnAlaLeuLeuValAspGluGluArgGly 65
 QY 401 AGCGTGTATGTTGGACCAAGATACATATTTTCATTCGACCTGGTTAATATC---AAG 457
 Db 66 ArgLeuPheValGlyAlaGluAlaHisValAlaSerLeuLeuAsnLeuAspAsnIleSerLys 85
 QY 458 GATTTTCAAGATGTTGGCCAGTATCTTACACCAGAGAGATGAATGCAATGGGCT 517
 Db 86 ArgAlaLysLeuLeuAlaTrpProAlaProValGluTrpArgGluGluCysAsnTrpAla 105
 QY 518 GGAAAGACATCTCTGAAAGATGCTCTAATTTTCATCAAGGTACTTTAAGGCATATAATCAG 577
 Db 106 GlyLysAspIleGlyThrGluCysMetAsnPheValLysLeuLeuHisAlaLysAsnArg 125
 QY 578 ACTACTTGTAGCTGTGGACGGGCTTTTCATCCAAATTTGCACCTACATTTGAAAT 637
 Db 126 ThrHisLeuLeuAlaCysGlyThrGlyAlaPheHisProThrCysAlaPheValGluVal 145
 QY 638 GGACATCATCTCGAGGCAATATTTTAAAGCTGGAGAACTCATCTTTGAAACGGCGCT 697
 Db 145

Db 146 GlyHisArgAlaGluGluProValLeuArgLeuAspProGlyArgIleGluAspGlyLys 165
 QY 698 GGGAGAGTCCATATGACCTAAGCTGTGTACAGCATCCCTTTTAAATAGATGAGAAATTA 757
 Db 166 GlyLysSerProTyrAspProArgHisArgAlaAlaSerValLeuValGlyGluGluLeu 185
 QY 758 TACTCTGGAACCTGACAGCTGATTTTATGGGGGAGACTTTTGTCTATCTTCCGAACCTCTGGG 817
 Db 186 TyrSerGlyValAlaAlaAspLeuMetGlyArgAspPheThrIlePheArgSerLeuGly 205
 QY 818 CACCACCACCAATCAGCAGCAGCAGCTGATTTCCAGTGGCTCAATGATCCAAAGTTC 877
 Db 206 GlnArgProSerLeuArgThrGluProHisAspSerArgTrpLeuAsnGlnProLysPhe 225
 QY 878 ATTAGTCCACCTCATCTCAGAGAGTGCATCTCTGAAGATGACAAATGATATACATTTTC 937
 Db 226 ValLysValPheTrpIleProGluSerGluAsnProAspAspAspLysIleTyrPhePhe 245
 QY 938 TTCCGTGAAATGCATAGATGGAGAACACTCTCT---GGAAAGCTACTCACCTCAGATTA 994
 Db 246 PheArgGluThrAlaValGluAlaAlaProAlaLeuGlyArgLeuSerValSerArgVal 265
 QY 995 GGTCAATATCAAGAATGACTTTTGGAGGACACAGACTCTGCTGAATAAATGGACACA 1054
 Db 266 GlyGlnIleCysArgAsnAspValGlyGlnArgSerLeuValAsnLysTrpThrThr 285
 QY 1055 TTCTCCTCAAGCTCGTCTGATTTGCTCAGTGCAGGTCCAAATGGCATTTGACATCATTTT 1114
 Db 286 PheLeuLysAlaArgLeuValCysSerValProGlyValGluGly---AspThrHisPhe 304
 QY 1115 GATCACTGACAGATGATTCCTAATGAACCTTTAAAGATCTCTAAATAATCCAGTTGTATAT 1174
 Db 305 AspGlnLeuGlnAspValPheLeuLeuSerSerArgAspHisArgThrProLeuLeuTyr 324
 QY 1175 GGAGTGTTCAGACTTCAGTAAACATTTTCAAGGATCAGCGCTGTGTATGATGATGATG 1234
 Db 325 AlaValPheSerThrSerSerSerIlePheGlnGlySerAlaValCysValTyrSerMet 344
 QY 1235 AGTGATGTGAGAAGGGTCTCTTGGTCCATATGCCACAGGATGGACCCAACTATCAA 1294
 Db 345 AsnAspValArgAlaPheLeuGlyProPheAlaHisLysGluGlyProMetHisGln 364
 QY 1295 TGGTGCCTTATCAAGGAAGAGTCCCTTATCCAGGCCAGGAACTTTGTCGCCAGCAAAACA 1354
 Db 365 TrpValSerTyrGlnGlyArgValProTyrProArgProGlyMetCysProSerLysThr 384
 QY 1355 TTTGGTGTGTTTGTACTCTACAAAGCCTTCTGATGATGTTATTAACCTTTTCCAGAGAT 1414
 Db 385 PheGlyThrPheSerSerThrLysAspPheProAspValIleGlnPheAlaArgAsn 404
 QY 1415 CATCCAGCATGTACATCCAGTGTCTTCTATGAACATCCCAATAGTATGATCAAAACG 1474
 Db 405 HisProLeuMetTyrAsnSerValLeuProThrGlyGlyArgProLeuPheGlnVal 424
 QY 1475 GATGTAAATATCAATTTTACAAATTTGTCTAGACCGAGTGGATGCGAGAGATGGACAG 1534
 Db 425 GlyAlaAsnTyrThrPheThrGlnIleAlaAlaAspArgValAlaAlaAspGlyHis 444
 QY 1535 TATGATGTATGTTTTCGGAACAGATGTTGGACCGTCTTAAAGTAGTTTCAATTCCT 1594
 Db 445 TyrAspValLeuPheIleGlyThrAspValGlyThrValLeuLysValIleSerValPro 464
 QY 1595 AAGCAGACTTGGTATGATTAGAAAGAGTCTCTGCTGGAAGAAATGACAGTTTTCGGGAA 1654
 Db 465 LysGlySerArgProSerAlaGluGlyLeuLeuGluGluLeuHisValPheGluAsp 484
 QY 1655 CCGACTGCTATTTCAGCAATGGAGCTTTTCCACTAAGCAGCAACAACTATATTTGGTTCA 1714
 Db 485 SerAlaAlaValThrSerMetGlnIleSerSerLysArgHisGlnLeuTyrValAlaSer 504
 QY 1715 ACGCTGGGTTGCCAGCTCCCTTTTACCGGTGTGATATTTACGGGAAAGCGGTGCT 1774
 Db 505 ArgSerAlaAlaGlnIleAlaLeuHisArgCysAlaAlaHisGlyArgValCysThr 524


```

QY 1775 GAGTGTGCTGCTCCGAGACCCCTACTGCTGCTGGATGGTCTGCTGCTGCTGCTAT 1834
Db 525 GluCyseLeuAlaArgAspProTyrCyseAlaTrpAspGlyValAlaCysThrArgPhe 544
QY 1835 TTTCCACTGCAAGAGAGCGCACAGACGACAGATATAGAAATGGAGACCCACTGACT 1894
Db 545 GlnProSerAlaLysArgPheArgGlnAspValArgAsnGlyAspProSerThr 564
QY 1895 CACTGTTAGACTTACACCATGATATACCATGCGCACAGCCCTGAA----- 1942
Db 565 LeuCyseSer-----GlyAspSerSerArgProAlaLeuLeu 576
QY 1943 GAGAGATCATCTATGCTGCTAGAGATAGTACACATTTTGGATGAGTCCGAGTCG 2002
Db 577 GluHisLysValPheGlyValGluGlySerSerAlaPheLeuGluCysGluProArgSer 596
QY 2003 CAGAGAGCGCTGCTCTATTGGCAATTCAGAGCGGAAATGAAGAGCGAAAGAGATC 2062
Db 597 LeuGlnAlaArgValGluTrpThrPheGlnArgAlaGlyValThrAlaHisThrGlnVal 616
QY 2063 AGAGTGTGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTAGTGTCTACAACAG 2122
Db 617 LeuAlaGluGluArgThrGluArgThrAlaArgGlyLeuLeuLeuArgArgLeuArg 636
QY 2123 AAGGATTCAGGCATTTACCTCTGCCATGCGGTGGACATGGCTTCATACAACTCTTCTT 2182
Db 637 ArgAspSerGlyValTyrLeuCyseAlaValGluGlnGlyPheThrGlnProLeuArg 656
QY 2183 AAGGTAACCTCGAAGTTCATTCACACAGACGATTTTGGAGAACTTCTTCATAAGATGAT 2242
Db 657 ArgLeuSerLeuHisValLeuSerAlaThrGlnAlaGluArgLeu----- 671
QY 2243 GATGGAGATGGCTCTTAAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAGAGTTC 2302
Db 672 -----AlaArgAlaGluGluAlaAlaProAlaAlaProGlyProGlySerLeu 687
QY 2303 TGGTACAGACTTCATGCTGCTCATCAACACCCCACTCTCAACAGTGTGATGATGTC 2362
Db 688 TrpTyrArgAspPheLeuGlnLeuVal----- 696
QY 2363 TGTGAACAAAGTTTGGAAAGGAGCGACCAACCAACGTCGCAAGGCGGACATACCCCA 2422
Db 697 -----GluProGly-----GlyGly 701
QY 2423 GGAACAGTAAACAAATGGAAG-----CACTTA 2449
Db 702 GlySerAlaAsnSerLeuArgMetCysArgProGlnProAlaLeuGlnSerLeuProLeu 721
QY 2450 CAAGAAATTAAGAAAGTAGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2497
Db 722 GluSerArgArgLysGlyArgAsnArgArgThrHisAlaProGluProArgAlaGluArg 741
QY 2498 GCACCCAGAGT 2509
Db 742 GlyProArgSer 745

```

RESULT 11

```

SM3B_MOUSE
ID SM3B_MOUSE STANDARD; PRT; 748 AA.
AC Q6217;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3B precursor (Semaphorin A) (Sema A).
GN SEMA3B OR SEMA OR SEMA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NMRI; TISSUE=Embryo;

```

```

RX MEDLINE=95267431; PubMed=7748561;
RA Puschel A.W., Adams R.H., Betz H.;
RT "Murine semaphorin D/collapsin is a member of a diverse gene family
and creates domains inhibitory for axonal extension.";
RL Neuron 14:941-948 (1995).
CC -!- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS.
CC -!- SUBCELLULAR LOCATION: Secreted (by similarity).
CC -!- DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE EMBRYO. LOW
MODERATE LEVELS FROM THEN UNTIL BIRTH.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
EMBL; X85990; CAA59982.1; -.
DR PIR; I48744; I48744.
DR MGD; MGI:107561; Sema3b.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003599; IG.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR001627; Sema.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 748 SEMAPHORIN 3B.
FT DOMAIN 55 496 SEMA.
FT DOMAIN 561 659 IG-LIKE C2-TYPE.
FT DOMAIN 698 702 POLY-GLY.
FT DOMAIN 723 743 ARG-RICH (BASIC).
FT DISULFID 643 709 BY SIMILARITY.
FT CARBOHYD 82 82 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 124 124 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 748 AA; 82894 MW; 1866B7D2397C9305 CRC64;

```

Alignment Scores:

```

Pred. No.: 2.88e-156 Length: 748
Score: 2059.50 Matches: 400
Percent Similarity: 66.93% Conservative: 112
Best Local Similarity: 52.29% Mismatches: 184
Query Match: 42.06% Indels: 69
DB: 1 Gaps: 15

```

US-09-774-490-1 (1-2709) x SM3B_MOUSE (1-748)

```

QY 281 AATGTGCCAAGCTGAATATTCTTACAAAGAAATGTTGGAATCCAAACATGTGATCACT 340
Db 27 AsnLeuProArgLeuArgLeuSerPheGlnGluLeuGlnAlaArgHisGlyValArgThr 46
QY 341 TTCAATGGCTTGGCCCAACAGTCCAGTTATCATCTTCTTCTTGGATGAGAACGAGT 400
Db 47 PheArg---LeuGluArgThrCysCysTyrGluAlaLeuLeuValAspGluGluArgGly 65
QY 401 AGGTGTATGTGGAGCAAGATCACAATATTTTCNTTCGACCTGGTTAATATC---AAG 457
Db 66 ArgLeuPheValGlyAlaGluAsnHisValAlaSerLeuSerLeuAsnIleSerLys 85
QY 458 GATTTCAAAAGATTGTGTGGCCAGTATCTTACACCAAGAGATGAATCAAGTGGGCT 517

```



```
Db      86  ArgAlaLysLeuValAlaTrpProAlaProValGluTrpArgGluGluCysAsnTrpAla 105
      ::::||||:|  |||||  |||:::|||||  |||||
Qy      518  GGAAAGACATCCTGAAAGATGTGCTAATTTTCATCAAGTACTTAAGGCATATATACAG 577
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      106  GlyLysAspIleGlyThrGluCysMetAsnPheValArgLeuLeuHisAlaIysrAsnHis 125
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      578  ACTCACTGTAGCCTGTGGAAAGCGGGGCTTTTCATCCAAATTCG 622
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      126  ThrHisLeuAlaCysArgThrGlyAlaPheHisProThrCysAlaLeuTrpArgTrp 145
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      623  -----ACCTACATTGAAATTGGACATCATCTTCAGGACACATATTTTAAG 667
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      146  AlaThrAlaGlyClyThrHisAlaSerThrGly 158
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      668  CTGGAGAACTCACAATTTTGAAGCGCGTGGGAAGAGTCCATATGACCTAAGCTGCTG 727
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      159  -----LysLeuGluAspGlyLysGlyLysThrProTyrAspProArgHisArg 174
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      728  ACAGCATCCCTTTTAATAGATGGAGAAATATATCTCTGGAACTGCGAGCTGATTTATGGG 787
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      175  ProProSerValLeuValGlyGluLeuTyrSerGlyValThrAlaAspLeuMetGly 194
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      788  CGAGACTTGTCTATCTCCGAACCTCTGGGCCACCCACCAATCAGGACAGCAGCAT 847
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      195  ArgAspPheThrIlePheArgSerLeuGlyGlnAsnProSerLeuArgThrGluProHis 214
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      848  GATTCAGGTGCTCAATGATCCAAAGTTTCATTAGTGGCCCACTCATCTCAGAGAGTGAC 907
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      215  AspSerArgTrpLeuAsnGluProLysPheValLysValPheTrpIleProGluSerGlu 234
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      908  AATCCTGAAGATGACAAATATATCTTTTCCTCGTGAATAATGCAATAGATGAGAAACAC 967
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      235  AsnProAspAspLysIleTyrPhePheArgGluSerAlaValGluAlaPro 254
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      968  TCT---GGAAAGCTACTCACCTAGATAGTATGATGATGATGATGATGATGATGATGAT 1024
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      255  AlaMetGlyArgMetSerValSerArgValGlyGlnIleCysArgAsnAspLeuGly 274
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      1025  CACGAAGTCTGGTGAATAAATGGCAACATTCCTCAAGCTCGTCTGATTTGCTCAGTG 1084
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      275  GlnArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeuValCysSerVal 294
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      1085  CCAGTCCAAATGGCATTGACACTCATTTTGTATGAACTCAGAGTGTATTCCTAATGAAC 1144
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      295  ProGlyValGluGly---AspThrHisPheAspGlnLeuGlnAspValPheLeuLeuSer 313
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      1145  TTTAAAGATCCTAAATCCAGTGTGTATGAGTGTGTATGAGTGTGTATGAGTGTGTATG 1204
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      314  SerArgAspArgGlnThrProLeuLeuTyrAlaValPheSerThrSerSerGlyValPhe 333
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      1205  AAGGGATCAGCCGTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1264
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      334  GlnGlySerAlaValCysValTyrSerMetAsnAspValArgAlaPheLeuGlyPro 353
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      1265  TATGCCCAAGGATGGACCCCACTCAATGGTGCCTTATCAAGGAGAGTCCCTTAT 1324
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      354  LeuProHisLysGluGlyProThrHisGlnTrpValSerTyrGlnGlyArgValProTyr 373
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      1325  CCACGGCCAGGAACCTTGTCCAGCAAAACATTTGTTGTTGTTGTTGTTGTTGTTGTTG 1384
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      374  ProArgProGlyMetCysProSerLysThrPheGlyThrPheSerSerThrLysAspPhe 393
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      1385  CCTGATGATGTTTAACTTTTCAAGAGTCTATCCAGCCATGTACATCACTCACTGTTCT 1444
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      394  ProAspAspValIleGlnPheGlyArgAsnHisProLeuMetCysTrpAsnProValLeuPro 413
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      1445  ATGAACAATCGCCCAATAGTATGATCAAAACGGATGTAATTTATCAATTTACACAAATGTC 1504
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      414  MetGlyGlyArgProLeuPheLeuGlnValGlyAlaGlyTyrThrPheThrGlnIleAla 433
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      1505  GTAGACCGAGTGGATGAGCAAGATGAGATGATGATGATGATGATGATGATGATGATGAT 1564
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```

RESULT 12

SM3E_HUMAN

ID SM3E_HUMAN

STANDARD;

PRT; 775 AA.

```
Db      434  AlaAspArgValAlaAlaAlaAspGlyHisTyrAspValLeuPheIleGlyThrAspVal 453
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      1565  GGGACCGTCTTAAAGTAGTTTCAATCTCAAGGAGACTTGGTATGATTATTAGAAGAGTTT 1624
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      454  GlyThrValLeuLysValIleSerValProLysGlyArgArgProAsnSerGluGlyLeu 473
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      1625  CTGCTGGAAAGAAATGACAGATTTTTCGGGAACCGACTGCTATTTCAGCAATGAGCTTCC 1684
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      474  LeuLeuGluGluLeuValPheGluAspSerAlaAlaIleThrSerMetGlnIleSer 493
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      1685  ACTAGCAGCAACAACTATATATGTTCAACGGCTGGGTGGCCAGCTCCCTTTACAC 1744
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      494  SerLysArgGlnGlnLeuTyrValAlaSerArgAlaAlaValAlaGlnIleAlaLeuHis 513
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      1745  CGGTGTGATATTTACGGGAACCGTGTCTGAGTGTCTGCTCGCCGAGACCTTACTGT 1804
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      514  ArgCysThrAlaLeuGlyArgAlaCysAlaGluCysCysLeuAlaArgAspProTyrCys 533
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      1805  GCTTGGGATGTTCTGCGATGTTCTCGCTATTTTCCACTGCAAGAGACGACAGACGA 1864
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      534  AlaTrpAspGlySerAlaCysThrArgPheGlnProThrAlaLysArgArgPheArg 553
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      1865  CAAGATATAAGAAATGGAGACCCACTGACTCACTGTTTCCAGACTTACCATGATATAC 1924
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      554  GlnAspIleArgAsnGlyAspProSerThrLeuCysSerGly 570
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      1925  CATGCCACAGCCCTGAAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGACATTTTG 1984
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      571  HisSerValLeuLeuGluLysLysValLeu---GlyValGluSerGlySerAlaPheLeu 589
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      1985  GAATGCACTCGAAGTCCGACAGAGCGCTGTCTATTGGCAATTCAGAGCGGGAATGAA 2044
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      590  GluCysGluProArgSerLeuGlnAlaHisValGlnTrpThrPheGlnGlyAlaGlyGlu 609
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      2045  GAGCGAAAGAGAGATCAGAGTGCATCATATCATCAGCAGCAGATCAGAGCCCTCTG 2104
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      610  AlaAlaHisThrGlnValLeuAlaGluGluArgValGluArgThrAlaArgGlyLeuLeu 629
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      2105  CTACGTAGCTTACAAACAGAGGATTCAGGCAATTTACCTCTGCGCATCGGTGGAAATGG 2164
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      630  LeuArgGlyLeuArgArgGlnAspSerGlyValTyrLeuCysValAlaValGluGlnGly 649
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      2165  TTCATACAACTCTTCTTAAGGTAACTCGTGAAGTCACTTGCACAGACATTTTGAAGAA 2224
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      650  PheSerGlnProLeuArgArgLeuValLeuHisValLeuSerAlaAlaGlnAlaGluArg 669
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      2225  CTTCTTCTAATAGATGATGATGAGTGCCTCTAAGACCAAGAAATGCTCCATAGCATG 2284
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      670  Leu-----AlaArgAlaGluGluAlaAlaProAla 680
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      2285  ACACCTAGCCAGAAGTCTGTTACAGACTTTCATGAGCTCATC-----AACCAC 2335
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      681  ProProGlyProLysLeuTyrTyrArgAspPheLeuGlnLeuValGluProGlyGlyGly 700
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      2336  CCCAATCTCAACACAGATGATGATGTTCTGTGAACAAGTTTGGAAAGGACCGCAAAACA 2395
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      701  GlyGlyAlaAsnSerLeu-----ArgMet 708
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      2396  CGTGGCAAGGCGGACGACATACCCAGGGACAGTAACAATGGAAGCAC-----2446
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      709  CysArgProGlnProGlyHis-----HisSerValAla 719
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      2447  TTACAAGAAATAAGAAAGGTAGAAACAGGAGGACCCAC-----GAATTTGAG 2494
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      720  AlaAspSerArgArgLysGlyArgAsnArgMetHisValSerGluLeuArgAlaGlu 739
      ::::||||:|  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy      2495  AGGGCACCCAGGAGT 2509
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db      740  ArgGlyProArgSer 744
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```


QY 1235 AGTGATGTGAGAGGGTGTCTTGGTCCATATGCGCCACAGGATGGACCCAACTATCAA 1294
 Db 347 SerSerIleArgAlaAlaPheAenGlyProTyAlaHisLeuGluGlyProGluTyHis 366
 QY 1295 TGGGTGCTTATCAAGGAAGAGTCCCTATCCACGGCCAGGAACCTTGTCACCAAGAAACA 1354
 Db 367 TrpSerValTyGluGlyValProTyProArgProGlySerCysAlaSerIleVal 386
 QY 1355 TTGTGTGGT---TTTGACTCTCAAGAGACCTTCTCGATGATGTTATTAACCTTTCAAGA 1411
 Db 387 AsnGlyGlyArgTyGlyThrThrIleAspTyProAspAlaIleArgPheAlaArg 406
 QY 1412 AGTCATCCAGCATGATCAATCCAGTGTCTTATGAACAATCGCCCAATAGTATCAA 1471
 Db 407 SerHisProLeuMetTyGlnAlaIleTyProAlaHisLeuValProIleLeuValIys 426
 QY 1472 ACCGATGTAATATCAATTTACACAAATGTGCTAGACCGAGTGGATGCGAAGATGGA 1531
 Db 427 ThrAspGlyLeuTyAsnLeuLysGlnIleAlaValAspArgValGluAlaGluAspGly 446
 QY 1532 CAGTATGATGTTATGTTATCGGAACAGATGTTGGACCGTCTCTTAAAGTAGTCTCAATT 1591
 Db 447 GlnTyAspValLeuPheIleGlyThrAspAsnGlyIleValLeuLysValIleThrIle 466
 QY 1592 CCTAAGGAGACTGTGATGATTTAGAGAGTCTGCTGGAAGAATGACAGTCTTCGG 1651
 Db 467 TyrAsnGlnGluMetGluSerMetGluGluValIleLeuGluGluLeuGlnIlePheIys 486
 QY 1652 GAACGAGCTGCTATTTCAGCAATGAGCTTCCACTAAGCAGCAACAATATATATCGT 1711
 Db 487 AspProValProIleIleSerMetGluIleSerIleValAspArgValGlnGlnLeuTyIleGly 506
 QY 1712 TCACCGCTGGGTGCTCCAGCTCCCTTTACACCGGTGTGATTTATTCGGGAACCGTGT 1771
 Db 507 SerAlaSerAlaValAlaGlnValArgPheHisHisCysAspMetTyGlySerAlaCys 526
 QY 1772 GCTAGTGTGCTGCTCCGCGAGACCTTACTGTGCTGCTGGATGTTGTCGATGTTCTCGC 1831
 Db 527 AlaAspCysCysLeuAlaArgAspProTyCysAlaIleAspGlyIleSerCysSerArg 546
 QY 1832 TATTTTCCACT-----GCAAGAGAGCGCACAGACGACAGATATAGAATGGA 1882
 Db 547 TyrTyProThrGlyThrHisAlaLysArgPheArgGlnAspValArgHisGly 566
 QY 1883 GACCCACTGACTCACTGTTACAGCTTACACCATGATAATCACCTGGCCAGCCCTGAA 1942
 Db 567 AsnAlaAlaGlnGlnCysPheGlyGlnGlnPheValGlyAspAlaLeuAspIleThrGlu 586
 QY 1943 GAGAAATCATCTATGCTGTAGAGATATAGTACACATTTTGGATGAGTCCGAGTCG 2002
 Db 587 GluHisLeuAlaTyGlyIleGluAsnAsnSerThrLeuLeuGluCysThrProArgSer 606
 QY 2003 CAGAGAGCGTGTCTATTGCAATCCAGAGCGCAATGAGACGGAAGAGAGATC 2062
 Db 607 LeuGlnAlaLysValIleTrpPheValGlnLysGlyArgGluThrArgLysGluGluVal 626
 QY 2063 AGATGGATGATCATATCATCAGACAGATCAAGGCTTCTGCTACGTAGTCTACAACAG 2122
 Db 627 LysThrAspAspArgValValLysMetAspLeuGlyLeuLeuPheLeuArgLeuHisLys 646
 QY 2123 AAGATTTCAGCAATATCTCTGCAATCGCGTGGAACTGGGTTCATCAAACTCTCTT 2182
 Db 647 SerAspAlaGlyThrTyPheCysGlnThrValGluHisSerPheValHisThrValArg 666
 QY 2183 AAGTAAACCTGGAATCATTCACACAGCATTTGGAGAACTTCTTCATAAGATGAT 2242
 Db 667 LysIleThrLeuGluValValGluGluGluLysValGluAspMetPheAsnLysAspAsp 686
 QY 2243 GATGAGATGGCTCTAAGACCAAA-----GAAATGTCTCAATAGCATCAACCTAGCCAG 2296
 Db 687 GluGluAspArgHisHisArgMetProCysProAlaGlnSerIleSerGlnGlyAla 706
 QY 2297 AAGTCTGTGACAGAGCTTATGATGAGTCTATCAACACCCCAATCTCAACACGATGAT 2356

Db 707 LysProTrpTyLysGluPheLeuGlnLeuIleGlyTySerAsnPheGlnArgValGlu 726
 QY 2357 GAGTCTCTGTGAACAAAGTTTGGAAAAAGGACCAACAAACAACTCGGCAAGGCCAGGACAT 2416
 Db 727 GluTyCysGluLysValTrpCysThrAspArgLysArgLysLeuLysMet----- 744
 QY 2417 ACCCCAGGAAAGTAAATGGAAGCATTACAA-----GAAAAATAAGAAAGGTAGAAAC 2473
 Db 745 -----SerProSerLysTrpLysTyAlaAsnProGlnGluLysLeuLysArgSer 761
 QY 2474 AGGAGGACCCAGCAATTTGAGAGGCCACCCAGG 2506
 Db 762 LysProGluHis-----TyrArgLeuProArg 770
 RESULT 13
 SM3E MOUSE STANDARD; PRT; 775 AA.
 AC P70275; O09078; O09079;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 3E precursor (Semaphorin H) (Sema H).
 GN SEMA3E OR SEMAH OR SEMH.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BAHB/C;
 RX MEDLINE=98175564; PubMed=9515811;
 RA Christensen C.R.L.; Klingelhoefer J.; Tarabykina S.; Hulgaard E.F.;
 RA Kramarov D.; Lukanidin E.;
 RT "Transcription of a novel mouse semaphorin gene, M-semah, correlates
 RT with the metastatic ability of mouse tumor cell lines.";
 RL Cancer Res. 58:1238-1244(1998).
 RN [2]
 RP REVISIONS.
 RA Christensen C.R.L.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: MAY BE INVOLVED IN EMBRYONIC DEVELOPMENT.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: DEVELOPING LUNGS, DEVELOPING SKELETAL
 CC ELEMENTS, AND VENTRAL HORNS OF THE DEVELOPING NEURAL TUBE.
 CC CORRELATES POSITIVELY WITH TUMOR PROGRESSION.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch)

 EMBL; Z80941; CAB02590.1; -;
 EMBL; Z93947; CAB07987.1; ALT_SEQ.
 EMBL; Z93948; CAB07988.1; ALT_SEQ.
 MGI; MGI:1340034; Sema3e.
 InterPro; IPR007110; Ig-like.
 InterPro; IPR003599; Ig.
 InterPro; IPR003006; Ig_MHC.
 InterPro; IPR003659; Plexin-like.
 InterPro; IPR001627; Sema.
 Pfam; PF00047; Ig; 1.
 Pfam; PF01403; Sema; 1.
 SMART; SM00409; IG; 1.
 SMART; SM00423; PSI; 1.
 SMART; SM00630; Sema; 1.
 PROSITE; PS50835; IG_LIKE; 1.

KW Signal: Immunoglobulin domain; Multigene family; Neurogenesis;

KW Developmental protein; Glycoprotein.

FT SIGNAL 1 25 POTENTIAL.

FT CHAIN 26 775 SEMAPHORIN 3E.

FT DOMAIN 241 540 SEMA.

FT DOMAIN 581 669 IG-LIKE C2-TYPE.

FT DOMAIN 737 770 ARG/LYS-RICH (BASIC).

FT DISULFID 654 729 BY SIMILARITY.

FT CARBOHYD 44 44 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 175 175 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 330 330 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 596 596 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 775 AA; 89503 MW; BA1690093D28F40 CRC64;

Alignment Scores:

Pred. No.: 4,12e-153 Length: 775

Score: 2020.00 Matches: 367

Percent Similarity: 67.62% Conservative: 155

Best Local Similarity: 47.54% Mismatches: 222

Query Match: 41.26% Indels: 18

DB: 1 Gaps: 8

US-09-774-490-1 (1-2709) x SM3E_MOUSE (1-775)

QY 218 ATTGCTGCTTTCTGGGAGTATTACTTACGACGAGCAAACTATCAGATGGGAG 277

DB 8 LeuThrLeuLeuLeuProGlyHisLeuLeuGluLeuThrProGlyHisSerAlaAsn 27

QY 278 AACAAATGTCAGGCTGAAATATCTCAAAAGAAATGTTGGAATCCAACTATGATC 337

DB 28 ProSerTyrProArgLeuArgLeuSerHisLysGluLeuLeuGluLeuAsnArgThrSer 47

QY 338 ACTTCAATGGCTGGCCAAACAGCTCCAGTTATCATCTTCTTTGATGAGAACGG 397

DB 48 IlePheGlnSerProLeuGlyPheLeuAspLeuHisThrMetLeuLeuAspGluTyrGln 67

QY 398 AGTAGCTGTATGTCGACCAAGATCATATTTTCACTTCGACTGTTAATATCAAG 457

DB 68 GluArgLeuValGlyGlyArgLeuValTyrSerLeuAsnGluArgValSer 87

QY 458 GAT---TTTCAAAAGATTCTGGCCAGTATCTTACACGACGAGATGAATGCAAGTGG 514

DB 88 AspGlyTyrArgGluIleTyrProSerThrAlaValLysValGluCysIleMet 107

QY 515 GCTGGAAGAGATCCTCGAAGATGTGCTAATTTTCATCAAGGTACTTAAGGCATATAT 574

DB 108 LysGlyLysAsp---AlaAsnGluCysAlaAsnTyrIleArgValLeuHisHisTyrAsn 126

QY 575 CAGACTCAGTCTGACGCTGGAGGGGCTTTTCATCCCAATTTGCACTTACATTGAA 634

DB 127 ArgThrHisLeuLeuThrCysAlaThrGlyAlaPheAspProHisCysAlaPheIleArg 146

QY 635 ATTGACATCATCTCGAGACCAATATTTTAAAGCTGGAGAACTCATATTTGAAAACGGC 694

DB 147 ValGlyHisHisSerGluGluProLeuPheHisLeuGluSerHisArgSerGluArgGly 166

QY 695 CQTGGAGAGATGCTCATATGACCTTAAGCTGCTGACAGCATCCTTTTAAATAGATGAGAA 754

DB 167 ArgGlyArgCysProPheAspProAsnSerSerPheValSerThrLeuValGlyAsnGlu 186

QY 755 TTATCTCTGGACTGCACCTGATTTTATGGGGCGAGACTTGTCTATCTTCCGACTCTT 814

DB 187 LeuPheAlaGlyLeuTyrSerAspTyrTrpGlyArgAspSerAlaIlePheArgSerMet 206

QY 815 GGGCACCAACCCCAATCAGACGACGACGATGATTTCCAGGTGGCTCAATGATCAAAAG 874

DB 207 GlyLysLeuGlyHisIleArgThrGluHisAspAspGluArgLeuLeuLysGluProLys 226

QY 875 TTCATTAGTCCCACTCATCTCAGAGAGTGCATCAATCTCTGAAGATGACAAAGATATCTTT 934

DB 227 PheValGlySerTyrMetIleProAspAsnGluAspArgAspAsnLysMetTyrPhe 246

QY 935 TTCTTCGTGAAATGCAATAGATAGGAGACACACTCTGAAAAAGCTACTCAGCCTAGAATA 994

DB 247 PhePheThrGluLysAlaLeuGluAlaGluAsnAsnAlaHisThrIleTyrThrArgVal 266

QY 995 GGTGAGATATCAAGAATGACTTTTGGAGGGCACAGAGTCTGGTGAATAAATAATGGAACA 1054

DB 267 GlyArgLeuCysValAsnAspMetGlyGlyGlnArgIleLeuValAsnLysTrpSerThr 286

QY 1055 TTCTCAAGCTCTGCTGATTTGCTCAGTCGAGGTCCTCAATGCGCATTTGACACTCATTTT 1114

DB 287 PheLeuLysAlaArgLeuValCysSerValProGlyMetAsnGlyIleAspThrTyrPhe 306

QY 1115 GATCAACTGACGAGTATCTTAATGAACCTTTAAAGATCTTAAAAATCCAGATCTGATAT 1174

DB 307 AspGluLeuGluAspValPheLeuLeuProThrArgAspProLysAsnProValIlePhe 326

QY 1175 GGAGTGTTCACGACTTCCAGTAACATTTTCAAGGGATCAGCGTGTGTATGATATGATG 1234

DB 327 GlyLeuPheAsnThrThrSerAsnIlePheArgGlyHisAlaValCysValTyrHisMet 346

QY 1235 AGTGATGTGACAGGGTGTCTTCTGGTCCATATGCCACAGGGATGGACCAACTATCAA 1294

DB 347 SerSerIleArgGluAlaPheAsnGlyProTyrAlaHisLysGluGlyProGluTyrHis 366

QY 1295 TGGGTGCTTATCAAGAAAGAGTCCCTTATCCAGGCGCAGCAACTTGTCCAGCAAAAACA 1354

DB 367 TrpSerLeuTyrGluGlyLysValProTyrProArgProGlySerCysAlaSerLysVal 386

QY 1355 TTTCGTGCT---TTGACTCTCAAGAGACCTTCTCTGATGATGTTATTAACCTTTCCAGA 1411

DB 387 AsnGlyGlyLysTyrGlyThrThrLysAspTyrProAspAspAlaIleArgPheAlaArg 406

QY 1412 AGTCATCCAGCCATGATCAATCCAGTGTCTTCTATGAACAATCCGCCAATAGTATCAA 1471

DB 407 IleAspProLeuMetTyrGlnProIleLysProValHisLysLysProIleLeuValLys 426

QY 1472 ACGATGTAAATTTATCAATTTTACAAATTTCTGTAGACCGAGTGGATGCAAGATGGA 1531

DB 427 ThrAspGlyLysTyrAsnLeuArgGlnLeuAlaValAspArgValGluAlaGluAspGly 446

QY 1532 CAGTATGATGTTATGTTTATCGGAACAGATTTGGGACCGTCTTCTTAAAGTGTATTTCAAT 1591

DB 447 GlnTyrAspValLeuPheIleGlyThrAspThrGlyIleValLeuLysValIleThrIle 466

QY 1592 CTAAGGAGACTTGGTATGATTTAGAGAGCTTCTGCTGGAGAAATACACAGTTTTCGG 1651

DB 467 TyrAsnGlnGluThrGluTrpMetGluGluValIleLeuGluGluLeuGluIlePheLys 486

QY 1652 GAACCGACTGTATTTTTCAGCAATGGAGCTTTCCACTAAGCAGCAACAACTATATATCGT 1711

DB 487 AspProAlaProIleIleSerMetGluIleSerSerLysArgGlnGlnLeuTyrIleGly 506

QY 1712 TCAACGGCTGGGTGGCCAGCTTCCCTTTACCGGTGTGATATTTACGGGAACCGGTGT 1771

DB 507 SerAlaSerAlaValAlaGlnValArgPheHisCysAspMetTyrGlySerAlaCys 526

QY 1772 CQTGAGTGTTCCTCGCCGAGACCTTACTGCTGCTGGATGTTCTGCAATTTCTCGC 1831

DB 527 AlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyIleSerCysSerArg 546

QY 1832 TATTTTCCACT-----GCAAGAGACGCAACAGCAACAGATATATGAAGATGGA 1882

DB 547 TyrTyrProThrGlyAlaHisAlaLysArgArgPheArgArgGlnAspValArgHisGly 566

QY 1883 GACCCACTGACTCTGTTTCAGACTTACCATGATATCATCCATGGCCACAGCCCTGAA 1942

DB 567 AsnAlaAlaGlnGlnCysPheGlyGlnPheValGlyAspAlaLeuAspArgThrGlu 586

QY 1943 GAGAGATCATCTATGCTGTAGAGAAATAGTAGACATTTTGGATTCAGTCCGAGTCG 2002

DB 587 GluArgLeuAlaTyrGlyIleGluSerAsnSerThrLeuLeuGluCysThrProArgSer 606

QY 2003 CAGAGAGCGCTGGTCTATTGGCAATTTCCAGAGGCGAAATGAAGAGCGAAAAAGAGATC 2062

```

Db 607 LeuGlnAlaLysValIleTrpPheValGlnLysGlyArgAspValArgLysGluGluVal 626
QY 2063 AGAGTGATGATCATATCATCATGAGCAGACAGGCTTCTGCTACGTAGTGTACAAACAG 2122
Db 627 LysThrAspArgValValLysMetAspLeuGlyLeuPheLeuArgValArgLys 646
QY 2123 AGGATTACGGAATTAATCTCTGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2182
Db 647 SerAspAlaGlyThrTyrPheCysGlnThrValGluHisAsnPheValHisThrValArg 666
QY 2183 AAGGTAACCTGGAAGTCTATGACACAGACGATTTGGAGAACTTCTTCAATAGATGAT 2242
Db 667 LysIleThrLeuGluValValGluGluHisLysValGluGlyMetPheHisLysAspHis 686
QY 2243 GATGAGATGGCTCTAAGACCAAGAAATGTCC-----AATAGCATGACACCT 2290
Db 687 GluGluGlu-----ArgHisLysMetProCysProLeuSerGlyMetSerGln 704
QY 2291 AGCCAGAGGTCTGTGACAGACATTCATGAGTTCATCAACACCCCAATCTCAACAG 2350
Db 705 GlyThrLysProTrpTyrLysGluPheLeuGlnLeuIleGlyTyrSerAsnPheGlnArg 724
QY 2351 ATGGATGATCTGTGAACAAGTTTGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2410
Db 725 ValGluGluTyrCysGluLysValTrpCysThrAspLysLysLysLysLysMet 744
QY 2411 GGACATACCCGAGGAGTAAACAAATGGAAGCACTTACAAAGAAATGGAAGGAGTGA 2470
Db 745 -----SerProSerLysTrpLysTyrAlaAsnProGlnGluLys---Arg 758
QY 2471 AACAGAGAGCCAGCAATTTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2506
Db 759 LeuArgSerLysAlaGluHisPheArgLeuProArg 770

```

RESULT 14

```

SM3E CHICK STANDARD; PRT; 785 AA.
AC 042237; Q90666;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Semaphorin 3E precursor (Collapsin-5) (COLL-5).
GN SEMA3E OR COLL5.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=97470887; PubMed=9331347;
RA Feiner L., Koppel A.M., Kobayashi H., Raper J.A.;
RT "Secreted chick semaphorins bind recombinant neuropilin with similar
RT affinities but bind different subsets of neurons in situ."
RL Neuron 19:539-545(1997).
RN [2]
RP SEQUENCE OF 244-543 FROM N.A.
RX MEDLINE=95329269; PubMed=7605628;
RA Luo Y., Shepherd I., Li J., Renzi M.J., Chang S., Raper J.A.;
RT "A family of molecules related to collapsin in the embryonic chick
RT nervous system."
RL Neuron 14:1131-1140(1995).
CC -!- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
CC CONES. COULD POTENTIALLY ACT AS REPULSIVE CUES TOWARD SPECIFIC
CC NEURONAL POPULATIONS. BINDS TO NEUROPILIN.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: COLLAPIN-1, -2, -3, AND -5 BIND TO
CC OVERLAPPING BUT DISTINCT AXON TRACTS.
CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
CC THIRD OF THE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.

```

```

CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC
CC This SMISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF022947; AAB0952.1; -.
CC EMBL; U28243; AAB6899.1; -.
CC InterPro; IPR003659; Plexin-like.
CC Pfam; PF01403; Sema; 1.
CC SMART; SM00423; PSI; 1.
CC SMART; SM00630; Sema; 1.
CC Signal; immunoglobulin domain; Multigene family; Neurogenesis;
CC Developmental protein; Glycoprotein.
CC SIGNAL 1 25 POTENTIAL.
CC CHAIN 26 785 SEMAPHORIN 3E.
CC DOMAIN 244 544 SEMA
CC DOMAIN 551 740 IG-LIKE C2-TYPE.
CC DISULFID 658 733 ARG/LYS-RICH (BASIC).
CC CARBOHYD 48 48 BY SIMILARITY.
CC CARBOHYD 130 130 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 500 500 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CONFLICT 246 246 N -> D (IN REF. 2).
CC CONFLICT 248 248 V -> I (IN REF. 2).
CC CONFLICT 250 250 L -> F (IN REF. 2).
CC SEQUENCE 785 AA; 90978 MW; E551BBF717630632 CRC64;

```

Alignment Scores:

Pred. No.:	5.9e-151	Length:	785
Score:	1993.00	Matches:	367
Percent Similarity:	67.26%	Conservative:	161
Best Local Similarity:	46.75%	Mismatches:	223
Query Match:	40.71%	Indels:	34
DB:	1	Gaps:	10

US-09-774-490-1 (1-2709) x SM3E_CHICK (1-785)

```

QY 218 ATGTGCTGCTTTCTTGGGGAGTATTACTACAGCAAGCAAACTATCAGATGGGAAG 277
Db 12 LeuLeuAlaLeuCysGlyLeuLeuLeuGluLeuProAlaGlyTyrHisAlaThrAsp 31
QY 278 AACAAATGCGCAAGGCTGAAATATTCTTACAAAGAAATGTTGGAAATCCCAATGTGATC 337
Db 32 ThrArgGlnProArgLeuArgLeuSerHisLysGluLeuTrpAspLeuAsnArgThrSer 51
QY 338 ACTTCAATGCTTGGCCACAGCTCCAGTTATCATATCTTCTTCTTGGATGAGGAACGG 397
Db 52 ValPheHisSerProPheGlyTyrLeuGlyLeuHisIleMetLeuLeuAspGluTyrGln 71
QY 398 AGTAGGCTGTATGTTGGCAAGCAAGCATCATATTTTTCATTCGACCTGGTTAAATATC--- 454
Db 72 GluArgLeuPheValGlyGlyArgAspLeuLeuTyrSerLeuSerLeuAspArgIleSer 91
QY 455 AAGGATTTTCAAAAGATTGTGTGCCAGTATCTTACACAGAGAGATGAATGCAAGTGG 514
Db 92 AsnAsnTyrArgGluIleHisTrpProSerThrProLeuGlnAlaGluGluCysIleIle 111
QY 515 GCTGAAAGACATCTCGAAGAAATGTCTTAATTTTCATCAAGGACTTAAGGCATATAAT 574
Db 112 LysGlyArgAsp---AlaAspGluCysAlaAsnTyrValArgValLeuHisArgTyrAsn 130
QY 575 CAGACTCACTGTACGCTGTGGAACGGGGCTTTTTCATCAATTCGACCTACATGAA 634
Db 131 ArgThrHisLeuLeuAlaCysGlyIleThrGlyAlaPheAspProValCysThrPheIleArg 150
QY 635 ATTGACATCATCTCTGAGGACAATATTTTAAAGCTGGAGAACTCACATTTTGAACACGGC 694

```

```

Db 151 ValGlyHisProSerGluAspHisLeuPheGlnLeuGluSerHisLysPheGluArgGly 170
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 695 CGTGGGAAGAGTCATATGACCCCTAAGCTGCTGACAGCATCCCTTTTAAATAGATGAGAA 754
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 171 ArgGlyArgCysProPheAspP-othrSerSerPheThrSerIleLeuGlyGlu 190
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 755 TTATACTCTGGAACCTGACCTGATTTTATGGGCGAGACTTTCCTATCTCCGAACTCTT 814
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 191 LeuPheThrGlyLeuTyrSerAspTyrTrpGlyArgAspAlaAlaValPheArgThrMet 210
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 815 GGGCACCCACCACCAATCAGGACAGACAGCATGATTCAGGTGGCTCAATGATCAAAAG 874
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 211 AsnArgMetAlaHisLeuArgThrGluProAspSerGluHisLeuLeuLysGluProLys 230
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 875 TTCATTAGTCCCACTCATCTCAGAGAGTGAACAATCCCTGAGATGACAAAGATATACTTT 934
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 231 PheValGlySerTyrMetIleProAspAsnGluAspHisAspAsnLysValTyrLeu 250
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 935 TTCTCCGCGTGAATCAATAGTAGAGACACTCTGGAAGAGCTACTCACCTAGATA 994
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 251 PhePheThrGluLysAlaLeuGluAlaGluThrSerThrHisAlaIleTyrThrArgVal 270
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 995 GGTCAATATGCAAGATCACTTTGGAGGCGACAGAAAGTCTGCTGAATAAATGGACAACA 1054
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 271 GlyArgValCysValAsnAspMetGlyGlyGlnArgIleValValAsnLysTrpSerThr 290
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1055 TTCTCAAGCTCGTCTGATTTGCTCAGTGCCAGGTCCAAATGGCAATGACACTCATTTT 1114
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 291 PheLeuLysThrArgLeuValCysSerValProGlyArgAsnGlyIleAspThrHisPhe 310
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1115 GATGAAGTCCAGATGATTTCTTAATGAATCTTAAAGATCCCTAAATCCAGTGTATAT 1174
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 311 AspGluLeuGluAspValPheLeuLeuGlnThrArgAspAsnLysAsnProValIlePhe 330
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1175 GGAGTGTTCACGACTTCCAGTAACTTTCAAGGATCAGCGGTGTGTATGATATGATGATG 1234
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 331 GlyLeuPheSerThrThrSerAsnIlePheArgGlyTyrAlaIleCysValTyrHisMet 350
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1235 AGTGATGTGAGAAGGTGTTCTTGTGTCATATGCCCCAGGATGGACCAACTATCAA 1294
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 351 AlaIleValArgAlaAlaPheAsnGlyProTyrAlaHisLysGluGlyProGluTyrTyr 370
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1295 TGGTGCTCTATCAAGAGAGTCCCTATCCAGCGCAGGAGACTGTGCCAGCAAAACA 1354
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 371 TrpAlaLeuTyrGluGlyLysValProTyrProArgProGlySerCysAlaSerLysVal 390
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1355 TTGTGTGGT---TTTGACTCTCAAAAGGACCTTCTCTGATGATGTTTATAACCTTTGCAAGA 1411
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 391 AsnGlyGlyLeuTyrThrThrThrLysAspTyrProAspGluAlaValHisPheAlaArg 410
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1412 AGTCATCCAGCCATGATCAATCCAGTGTGTTCTTATGAACAATCGCCCAATAGTATCAA 1471
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 411 SerHisProLeuMetTyrGlnProIleLysProValHisLysArgProIleLeuValLys 430
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1472 ACGGATGTAATATCAATTTACAAATTTGCTAGACCGAGTGGATCGACAGATGGA 1531
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 431 ThrAspGlyLysTyrAsnLeuLysGlnIleAlaValAspArgValGluAlaGluAspGly 450
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1532 CAGTATGATGTTATGTTTATCGGAACAGATGTTGGGACCGTTCCTTAAAGTAGTTTCAATT 1591
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 451 GlnTyrAspValLeuPheIleGlyThrAspAsnGlyIleValLeuLysValIleThrIle 470
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1592 CCTAAGGAGACTGGTATGATTTAGAGAGTGTCTGCTGGAAGAATAGACAGTTTTCGG 1651
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 471 TyrAsnGlnGluThrGluSerMetGluGluValIleLeuGluGluLeuGlnValPheLys 490
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1652 GAACCCAGTCTATTTTCAACATGGAGCTTCCACTTAACGACGACCAACTATATATGCT 1711
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 491 ValProIleProIleLeuSerMetGluIleSerLysArgGlnGlnLeuTyrIleGly 510
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1712 TCAACCGCTGGGTTCGCCAGTCCCTTTTACACCGGTGTGATATTTTACGGGAAGCGTGT 1771
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::

```

```

Db 511 ThrGluSerValIleAlaGlnValLysPheHisGlnCysAspMetTyrGlyThrAlaCys 530
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1772 GCTGAGTGTTCCTCGCCGAGACCCCTTACTGCTGCTGGGATGGTTCTGCATGTTCTCGC 1831
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 531 AlaAspCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlyIleSerCysSerArg 550
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1832 TATTTTCCACT---GCAAAGAGACGCAACAGACGACCAAGATATATAGAAATGGA 1882
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 551 TyrTyrProThrGlyMetGlnAlaLysArgPheArgArgGlnAspValArgHisGly 570
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1883 GACCCACTGACTCTGTTTACAGATTACCATGATATATCACCATGGCCACAGCCCTGAA 1942
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 571 AsnAlaAlaGlnGlnCysPheGlyGlnGlnPheIleGlyGluValLeuGluLysThrGlu 590
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 1943 GAGAAATCATCTATGTTGTTAGAGAAATAGTACCATTTTTTGGATGCGAGTCCGAGTCG 2002
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 591 GluArgLeuValTyrGlyIleGluTyrAsnSerThrLeuLeuGluTyrThrProArgThr 610
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 2003 CAGAGAGCGTGTCTATTGGCAATTCCAGAGGGGAAATGAAGAGCAAAAGAGAGATC 2062
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 611 LeuGlnAlaLysValAsnTrpPheValGlnArgAlaHisGluThrLysLysGluGluVal 630
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 2063 AGAGTGGATGATCATATCATCAGGACAGATCAAGGCTTCTGCTACGTAGTCTACAAACAG 2122
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 631 LysThrAspGluArgIleIleLysMetAspLeuGlyLeuLeuPheLeuLysLeuHisArg 650
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 2123 AAGATTTCAGCAATACCTCTGCTGCGGTGGAACATGGTTTCATACAACTCTTCTT 2182
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 651 LeuAspAlaGlyThrTyrPheCysGlnThrValGluHisSerIleValHisThrValArg 670
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 2183 AAGTTAAACCTCGGAAGTCATTGACACAGACGATTTGGAAGAACTTCTTATAAAGATCAT 2242
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 671 LysIleThrLeuGluIleValGluGluArgValAspGluMetPheSerLysAspTyr 690
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 2243 GATGGAGATGCTCTTAAGACCAAAAGAAATGCTCAATAGCATG---ACACCTAGCCAG--- 2296
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 691 Glu-----GluGluIleSerHisLysMetProCysProMetGlnSer 704
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 2297 -----AAGTCTGTTACAGAGACTTCATGTCAGCTCATCAACACACCC 2338
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 705 AsnIleProGlnValSerLysProTyrLysGluPheLeuGlnLeuIleGlyTyrSer 724
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 2339 AATCTCAACGATGGATGAGTCTCTGCAACAAGTTTGGAAAAGGCGCAAAACAAAGT 2398
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 725 AsnPheGlnArgValGluGluTyrCysGluLysValTyrCysThrAspLysLysArgLys 744
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 2399 CGGCAAGGCCAGGACATACCCAGGGAAACAGTAACAAATCGAAGCACTTA----- 2449
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 745 LysLeuLysMet-----SerProSerLysTrpLysTyrAlaAsnProGln 759
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 2450 -----CAAGAAATAAGAAAGGTAGAACAGAGGAGGCCACCAAGATTTCAGAGG 2497
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 760 GluLysArgGlnAspGlnGluLysLysAlaArgIleArgProGluHis-----TyrArg 777
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
QY 2498 GCACCCAGGAGTGTCT 2512
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 778 LeuProArgAsnIle 782
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::

RESULT 15
SM3F HUMAN STANDARD; PRT; 785 AA.
ID SM3F HUMAN Q13274; Q13372; Q15704;
AC Q13274; Q13274; Q13372; Q15704;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3F precursor (Semaphorin IV) (Sema IV) (Sema III/F).
GN SEMA3F.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

RC TISSUE=Petal brain;
RX MEDLINE=96226360; PubMed=8649831;
RA Roche J., Boldog F., Robinson M., Robinson L., Varella-Garcia M.,
RA Swanton M., Maggoner B., Fishel R., Franklin W., Gemmill R.,
RA Drabkin H.;
RT "Distinct 3p21.3 deletions in lung cancer and identification of a new
RT human semaphorin.";
RL Oncogene 12:1289-1297(1996).
RN [2]
RP SEQUENCE FROM N.A., AND VARIANT MET-503.
RC TISSUE=Petal brain;
RX MEDLINE=96230324; PubMed=8786119;
RA Xiang R.-H., Hensel C.H., Garcia D.K., Carlson H.C., Kok K.,
RA Daly M.C., Kerbacher K., van den Berg A., Veldhuis P., Buys C.H.C.M.,
RA Naylor S.L.;
RT "Isolation of the human semaphorin III/F gene (SEMA3F) at chromosome
RT 3p21, a region deleted in lung cancer.";
RL Genomics 32:39-48(1996).
RN [3]
RP SEQUENCE FROM N.A.
RA Nelson J., Biewald T.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 394-436 FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=96210603; PubMed=8633026;
RA Sekido Y., Bader S., Latif F., Chen J.-Y., Duh P.-M., Wei M.-H.,
RA Albanesi J.P., Lee C.-C., Lerman M.I., Minna J.D.;
RT "Human semaphorins A(V) and IV reside in the 3p21.3 small cell lung
RT cancer deletion region and demonstrate distinct expression patterns.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:4120-4125(1996).
CC -!- FUNCTION: MAY PLAY A ROLE IN CELL MOTILITY AND CELL ADHESION.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY BUT DIFFERENTIALLY IN A
CC VARIETY OF NEURAL AND NONNEURAL TISSUES. THERE IS HIGH EXPRESSION
CC IN MAMMARY GLAND, KIDNEY, FETAL BRAIN, AND LUNG AND LOWER
CC EXPRESSION IN HEART AND LIVER.
CC -!- DEVELOPMENTAL STAGE: DETECTED AS EARLY AS EMBRYONIC DAY 10.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; U33920; AAC50568.1; -;
DR EMBL; U38276; AAB18276.1; -;
DR EMBL; AC000063; BAB46344.1; -;
DR EMBL; U32171; AAS06011.1; -;
DR EMBL; U32172; AAB06012.1; -;
DR Genew; HGNC:10728; SEMA3F.
DR MIM; 601124; -;
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR001627; Sema.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; Ig; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS50835; IG LIKE; 1.
KW Signal; Immunoglobulin domain; Multigene family; Glycoprotein;
KW Polymorphism. 1 18
FT SIGNAL 19 785
FT CHAIN

FT	DOMAIN	272	569	SEMA.
FT	DOMAIN	605	690	IG-LIKE C2-TYPE.
FT	DOMAIN	758	779	ARG/LYS-RICH (BASIC).
FT	DISULFID	678	746	BY SIMILARITY.
FT	CARBOHYD	53	53	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	126	126	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	474	474	A -> G (in dbSNP:1046955).
FT	VARIANT	503	503	/FTid=VAR_011820.
FT	VARIANT	503	503	L -> M.
FT	CONFLICT	153	183	/FTid=VAR_008855.
FT	CONFLICT	270	270	MISSING (IN REF. 2).
FT	CONFLICT	473	473	MISSING (IN REF. 2).
FT	CONFLICT	473	473	A -> S (IN REF. 2).
SQ	SEQUENCE	785 AA;	88381 MW;	FE3FC796EEC1608E CRC64;

Alignment Scores:
Pred. No.: 9,99e-147 Length: 785
Score: 1940.00 Matches: 372
Percent Similarity: 63.98% Conservative: 136
Best Local Similarity: 46.85% Mismatches: 226
Query Match: 39.62% Indels: 60
DB: 1 Gaps: 11

US-09-774-490-1 (1-2709) x SM3P_HUMAN (1-785)

QY	227	CTTTCTGGGAGTATTACTTTTAC	---GCAAGAGCAAACTAT	-----CAGAATGGGAG	277
DB	7	LeuLeuTrpAlaSerLeuLeuThrGlyAlaTrpProSerPheProThrGlnAspHisLeu	26		
QY	278	AACATGTGCGAAGCTGAAATTTCT	TACAAAGAAATGTTGGAATCCAC	ATGTGATC	337
DB	27	ProAlaThrProArgValArgLeuSerPheLysGluLeuLysAlaThrGlyThrAlaHis	46		
QY	338	ACTTTCAATGCTTGGCCACAGCT	CACAGTATCATACCTCTCTT	TTTGGATGAGAACGG	397
DB	47	PhePheAsnPheLeuLeuAsnThrAspTyrArgIleLeuLeuLysAspGluAspHis	66		
QY	398	AGTAGGCTGTATGTTGGAGCAAG	GATCACAATATTTTCATTCGAC	CTCGTTAATATCAAG	457
DB	67	AspArgMetTyrValGlySerLys	AspTyrValLeuSerLeuAspLeuHis	AspLeuAsn	86
QY	458	GATTTTCAAAAGATGTG	---TGCCAGTAGTCTTACACCA	GAGAGATGATGCAGTGG	514
DB	87	ArgGluProLeuIleIleHisTrp	AlaAlaSerProGlnArgIleGlu	GluCysValLeu	106
QY	515	GCTGAAAGACATCTCGAAGAA	TGCTCAATTTTCATCAAGGT	ACTTAAGGCATATAAT	574
DB	107	SerGlyLysAspValAsnGlyGlu	CysGlyAsnPheValArgLeuIleGln	ProTrpAsn	126
QY	575	CAGACTCACTTGTCGCTGTGGA	ACGGGCTTTTCATCCAAATTTG	CACCTACATTCGA	634
DB	127	ArgThrHisLeuTyrValCysGly	ThrGlyAlaTyrAsnProMetCys	ThrTyrValAsn	146
QY	635	ATTGACATCATCT	-----	-----	649
DB	147	ArgGlyArgAlaGlnAlaThr	ProTrpThrGlnThrGlnAlaVal	ArgGlyArgGly	166
QY	650	-----	-----	-----GAGGACAAATTT	661
DB	167	SerArgAlaThrAspGlyAlaLeu	ArgProMetProThrAlaProArg	GluAspTyrIle	186
QY	662	TTTAAGCTGGAGAACTCACAT	TTTGAACCGCGTGGAGAGAGT	TCCATATGACCCCTAAG	721
DB	187	PheTyrLeuGluProGluArgLeu	SerGlyLysGlyCysProTyrAsp	ProLys	206
QY	722	CTGCTGACAGATCCCTTTTAA	TAGATGGAGAAATTAATCT	TGGAACGTGACGTGATTTT	781
DB	207	LeuAspThrAlaSerAlaLeu	IleAsnGluLeuTyrAlaGly	ValTyrIleAspPhe	226
QY	782	ATGGGGCGAGACTTTGCTAT	CTTCCGAACCTTTGGGACAC	CACACCCCAATCAGACAGAG	841
DB	227	MetGlyThrAspAlaAlaIle	PheArgThrLeuGlyLysGln	ThrAlaMetArgThrAsp	246


```
QY 842 CAGCATGATTCAGGTGGCTCAATGATCCAAAGTTCTTAGTGGCCACCTCATCTCAGAG 901
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 247 GNTYRASNSerArgTrpLeuAsnAspProSerPheIleHisAlaGluLeuIleProAsp 266
QY 902 AGTGACAATCCTGAAGATGACAAAGTATATCTTTCTCCGGTGAATAATCAATAGATGGA 961
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 267 SerAlaGluArgAsnAspAspLysLeuTyrPhePheArgGluArgSerAlaGluAla 286
QY 962 GAACACTCTGGAAAAGCTACTCAGCTAGATAAGTGCATATGCAAGATGACTTTGGA 1021
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 287 ProGlnSer---ProAlaValTyrAlaArgIleGlyArgIleCysLeuAsnAspGly 305
QY 1022 GGGCAGAGAAGTCTGGTGAATAAATGGACAACATTCCTCAAGCTCGTGTGATTGGCTCA 1081
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 306 GlyHisCysCysLeuValAsnLysTrpSerThrPheLeuLysAlaArgLeuValCysSer 325
QY 1082 GTGCCAGGTCCAAATGGCACTGACACTCATTTTGTATGAAGTGCAGGATGATTTCTTAATG 1141
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 326 ValProGlyGluAspGlyIleGluThrHisPheAspGluLeuGlnAspValPheValGln 345
QY 1142 ACTTTAAAGATCCTAAATCCAGTTGTATATGAGGTGTTTACGACTTCCAGTAACTT 1201
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 346 GlnThrGlnAspValArgAsnProValIleTyrAlaValPheThrSerSerGlySerVal 365
QY 1202 TTCAAGGGATCAGCGTGTGTATGATACATGATGATGATGATGATGATGATGATGATGAT 1261
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 366 PheArgGlySerAlaValCysValTyrSerMetAlaAspIleArgMetValPheAsnGly 385
QY 1262 CCATATGCCACAGGATGGACCCCACTATCAATGGTGGTCCCTTATCAAGAAAGAGTCCCC 1321
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 386 ProPheAlaHisLysGluGlyProAsnTyrGlnTrpMetProPheSerGlyLysMetPro 405
QY 1322 TATCCAGGCCAGGAATCTGTCCACGCAAAACATTT---GGTGGTTTTCAGTCTCAAGAG 1378
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 406 TyrProArgProGlyThrCysProGlyGlyThrPheThrProSerMetLysSerThrLys 425
QY 1379 GACCTTCCTGATGATGATATACCTTTGCAAGAAGTCATCCAGCCATGTACAAATCCAGTG 1438
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 426 AspTyrProAspGluValIleAsnPheMetArgSerHisProLeuMetTyrGlnAlaVal 445
QY 1439 TTTCTTATGAACAATCGCCCAATAGTATGATCAAAACGGATGTAAATTTATCAATTTACAAA 1498
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 446 TyrProLeuGlnArgArgProLeuValValArgThrGlyAlaProTyrArgLeuThrThr 465
QY 1499 ATGTGCTGATAGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1558
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 466 IleAlaValAspGlnValAspAlaAlaAspGlyArgTyrGluValLeuPheLeuGlyThr 485
QY 1559 GATGTTGGGACCGTCTTAAAGTAGTTTCAATTTCTTAAGGAGACTTGGTATGATTTAGAA 1618
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 486 AspArgGlyThrValGlnLysValIleValLeuProLysAspAsp---GlnGluLeuGlu 504
QY 1619 GAGGTTCTGCTGGAGAAATGACAGTTTTTCGGGAACCGACTGCTATTTCAGCAATGGAG 1678
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 505 GluLeuMetLeuGluGluValGluValPheLysAspProAlaProValLysThrMetThr 524
QY 1679 CTTTCCACTAGCAGCAACAATATATATATGTTTCAACGGCTGGGGTCCCGCAGCTCCCT 1738
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 525 IleSerSerLysArgGlnGlnLeuTyrValAlaSerAlaValGlyValThrHisLeuSer 544
QY 1739 TTACACCGGTGTGATATTTACGGGAAGCGTGTGCTGAGTGTGCTCGCCCGAGACCCCT 1798
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 545 LeuHisArgCysGlnAlaTyrGlyAlaAlaCysAlaAspCysCysLeuAlaArgAspPro 564
QY 1799 TACTGTGCTGGGATGGTCTGTCATGTTCTCGCTATTTTCCCACTGCAAAAGAGACGACA 1858
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 565 TyrCysAlaTyrAspGlyGlnAlaCysSerArgTyrThrAlaSerSerLysArgArgSer 584
QY 1859 AGACGACAAGATATAGAATAATGGAGCCCACTGACTGCTGTTTACAGACTTACACCATGAT 1918
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 585 ArgArgGlnAspValArgHisGlyAsnProIleArgGlnCysArgGlyPhe-----Asn 602
QY 1919 AATCACCATGGCCACAGCCCTGAAGAGAGAGAAATCATCTATGTTGTAGAGAAATAGTAGCACA 1978
```

Search completed: August 3, 2003, 10:02:36
Job time : 70.5 secs

```
DB 603 SerAsnAlaAsnLysAsnAlaValGluSerValGlnTyrGlyValAlaGlySerAlaAla 622
QY 1979 TTTTGGAAATCGAGTCCGAAGTCGAGAGAGCGGTGCTATTGGCAATTTCCAGAGGCGA 2038
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 623 PheLeuGluCysGlnProArgSerProGlnAlaThrValLysTrpLeuPheGlnArgAsp 642
QY 2039 AATGAAGAGCGAAAGAAGAGATCAGATGAGTGCATATCATCATCAGACAGATCAAGGC 2098
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 643 ProGlyAspArgArgGluIleArgAlaGluAspArgPheLeuArgThrGlnGlnGly 662
QY 2099 CTTCTGCTTACCTAGTCTACAAACAGAAAGGATTCAGGCAATTTACCTCTGCCATCGGTGGAA 2158
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 663 LeuLeuLeuArgAlaLeuGlnLeuSerAspArgGlyLeuTyrSerCysThrAlaThrGlu 682
QY 2159 CATGGTTTCATACAAACTCTTCTTAAGGTAAACCTCGGAAAGTCATTGACACAGAGATTG 2218
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 683 AsnAsnPheLysHisValValThrArgValGlnLeuHisValLeuGlyArgAspAlaVal 702
QY 2219 GAAGAACTTCTTCATAAAGATGATGATGATGCTCTAAGACCACAAAGAAATGTCCAAT 2278
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 703 HisAlaAlaLeuPhe-----ProProLeuSerMet 712
QY 2279 AGCATGACACATAGCCAGAGAGTCTGG-----TACAGAGACTTTCATGCAG 2323
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 713 SerAlaProProProProGlyAlaGlyProProThrProProTyrGlnGluLeuAlaGln 732
QY 2324 CTCTATCAACCAACCCCAATCTCAACACAGATGATGATGATGATGATGATGATGATGATGAT 2383
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 733 LeuLeuAlaGlnProGluValGlyLeuIleHisGlnTyrCysGlnGlyTyrTrpArgHis 752
QY 2384 GACCGAAACAACTCGCGCAAGGCCAGGACATACCCCGAGGAAACAGTAACAAATGGAAG 2443
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 753 ValProProSerProArgGluAlaProGly-----AlaProArgSerPro 767
QY 2444 CACTTACAAGAAATAAGAAAGGTAGAAACAGGAGGACCCAC 2485
DB |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 768 GluProGlnAspGlnLysLysProArgAsnArgArgHisHis 781
```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 10:29:36 ; Search time 28 Seconds
(without alignments)
9099.667 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 895

Sequence: 1 aatctttttatctatcgatg.....aggcttttttctctaataacc 2709

Scoring table:

OLIGO Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 127863 seqs, 47026705 residues

Word size: 1

Total number of hits satisfying chosen parameters: 255550

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlh
-O=/cgn2.1/USPTO.spool/US09774490/runat_03082003.102927.4449/app_query.fasta_1.2887
-DB=Swissprot_41 -QFMT=fastan -SUFFIX=oligo.rsp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=100 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=75 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09774490 @CGN 1.1.33 @runat_03082003.102927.4449 -NCPU=3
-NO_MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	771	86.1	771	1 SM3A_HUMAN	Q14563 homo sapien
2	92	10.3	772	1 SM3A_MOUSE	O08665 mus musculus
3	92	10.3	772	1 SM3A_RAT	Q63548 rattus norv
4	58	6.5	772	1 SM3A_CHICK	Q90607 gallus gall
5	35	3.9	860	1 SM3A_BRARE	Q9w7j1 brachydanio
6	24	2.7	778	1 SM3B_BRARE	Q9w686 brachydanio
7	20	2.2	748	1 SM3B_MOUSE	Q62177 mus musculus
8	16	1.8	764	1 SM22_BRARE	Q9w696 brachydanio
9	15	1.7	749	1 SM3B_HUMAN	Q13214 homo sapien
10	15	1.7	751	1 SM3C_CHICK	O42236 gallus gall
11	15	1.7	751	1 SM3C_HUMAN	Q99985 homo sapien
12	15	1.7	751	1 SM3C_MOUSE	Q62181 mus musculus
13	14	1.6	777	1 SM3D_HUMAN	O95025 homo sapien
14	13	1.5	761	1 SM3D_CHICK	Q90663 gallus gall
15	13	1.5	775	1 SM3E_HUMAN	O15041 homo sapien
16	13	1.5	775	1 SM3E_MOUSE	P70275 mus musculus
17	13	1.5	785	1 SM3E_CHICK	O42237 gallus gall
18	13	1.5	785	1 SM3F_HUMAN	Q13275 homo sapien

19	13	1.5	785	1 SM3F_MOUSE	O08632 mus musculus
20	10	1.1	834	1 SM4C_MOUSE	Q64151 mus musculus
21	9	1.0	295	1 SM4D_CHICK	Q90665 gallus gall
c	22	1.0	688	1 C1S_HUMAN	P09871 homo sapien
23	9	1.0	782	1 SM4B_MOUSE	Q62179 mus musculus
24	9	1.0	832	1 SM4B_HUMAN	Q9npr2 homo sapien
25	9	1.0	861	1 SM4D_MOUSE	O09126 mus musculus
26	9	1.0	862	1 SM4D_HUMAN	Q92854 homo sapien
27	9	1.0	913	1 IF38_HUMAN	Q99613 homo sapien
28	9	1.0	1244	1 SLA1_YEAST	P32790 saccharomyc
c	29	1.0	4870	1 RYR3_HUMAN	Q15413 homo sapien
30	9	1.0	4967	1 RYR2_HUMAN	Q92736 homo sapien
c	31	1.0	4969	1 RYR2_RABIT	P30957 oryctolagus
32	8	0.9	50	1 PND1_ECOLI	P11902 escherichia
33	8	0.9	50	1 PND2_ECOLI	P16477 escherichia
c	34	0.9	75	1 MTRB_BACSU	P19466 bacillus su
35	8	0.9	257	1 VSPC_TRIGA	O13062 trimeresu
c	36	0.9	368	1 GRBB_BACSU	P39570 bacillus su
37	8	0.9	374	1 SPB8_HUMAN	P50452 homo sapien
38	8	0.9	377	1 SH1D_CANFA	P11614 canis fami
c	39	0.9	462	1 IBMP_SOCMV	P15628 soybean chl
40	8	0.9	479	1 PBP4_HAETN	P45161 h penicilli
41	8	0.9	579	1 DLD1_KLULA	Q12627 kluyveromyc
42	8	0.9	697	1 Y29C_SCHPO	O13773 schizosacch
43	8	0.9	760	1 SM4A_MOUSE	Q62178 mus musculus
44	8	0.9	761	1 SM4A_HUMAN	Q9h3s1 homo sapien
45	8	0.9	837	1 SM4G_MOUSE	Q9whu7 mus musculus
46	8	0.9	838	1 SM4G_HUMAN	Q9ntn9 homo sapien
47	8	0.9	966	1 F1B1_PETWA	P02674 petromyzon
48	8	0.9	1074	1 SM5A_HUMAN	Q13591 homo sapien
49	8	0.9	1077	1 SM5A_MOUSE	Q62217 mus musculus
c	50	0.9	1409	1 RPOC_PSEPU	P19176 pseudomonas
c	51	0.9	5035	1 RYR1_PIG	P16960 sus scrofa
c	52	0.9	5037	1 RYR1_RABIT	P11716 oryctolagus
c	53	0.9	5038	1 RYR1_HUMAN	P21817 homo sapien
54	7	0.8	37	1 PSAJ_EUGER	P30394 euglena gra
c	55	0.8	54	1 RSL4_PYRAE	Q8zvw1 pyrobaculum
c	56	0.8	63	1 YPH2_SYNPF	P32040 synethococ
57	7	0.8	68	1 YRP3_AZOCA	P33988 azorhizobiu
58	7	0.8	99	1 YL13_ARCFU	O28167 archaeoglob
59	7	0.8	100	1 KRP1_COLLI	Q9pr15 columba liv
60	7	0.8	100	1 KRP2_COLLI	O93499 columba liv
61	7	0.8	100	1 KRP3_COLLI	O93500 columba liv
62	7	0.8	100	1 KRP4_COLLI	Q9psv3 columba liv
c	63	0.8	102	1 YC54_AQUAE	Q67295 aquifex aeo
c	64	0.8	106	1 INS2_XENLA	P12707 xenopus lae
c	65	0.8	106	1 Y094_HAETN	P43939 haemophilus
c	66	0.8	118	1 LG22_RHOCA	P23460 rhodobacter
c	67	0.8	119	1 CRCB_NEIMA	Q9jull1 neisseria m
c	68	0.8	119	1 CRCB_NEIMB	Q9jz96 neisseria m
c	69	0.8	128	1 CRCB_PASMU	Q9ck17 pasteurella
c	70	0.8	133	1 ATP6_THUOB	Q36090 thunnus obe
c	71	0.8	134	1 SRP_CHLPS	P28164 chlamydia p
c	72	0.8	134	1 YBJ5_CANAL	P53716 candida alb
c	73	0.8	139	1 UGR1_MOUSE	Q920h1 mus musculus
c	74	0.8	141	1 HBA5_XENLA	P06638 xenopus lae
c	75	0.8	144	1 PSGB_HUMAN	Q9u774 homo sapien
c	76	0.8	153	1 PYR1_ARCFU	O30129 archaeoglob
c	77	0.8	153	1 PYR1_VIBPA	Q8c1f7 vibrio para
c	78	0.8	153	1 PYR1_VIBVU	P46491 haemophilus
c	79	0.8	160	1 CRCB_HAETN	Q9zds3 rickettsia
c	80	0.8	163	1 Y252_RICPR	Q9zds3 rickettsia
c	81	0.8	167	1 YC51_METUA	Q58649 methanococ
c	82	0.8	168	1 ZG42_XENLA	P18720 xenopus lae
c	83	0.8	170	1 PER2_HORVU	Q01548 hordeum vul
c	84	0.8	176	1 PTHR_CHICK	P17251 gallus gall
c	85	0.8	176	1 VGB_CVPE3	P33467 feline ente
c	86	0.8	185	1 EFP_THETN	Q8rae2 thermoanaer
c	87	0.8	188	1 Y816_AERPE	Q9ydv3 aeropyrum p
c	88	0.8	192	1 ARL7_HUMAN	P56559 homo sapien
c	89	0.8	193	1 NU2M_PARTE	P55777 paramecium
c	90	0.8	200	1 MODB_BPT4	P39421 bacterioph
c	91	0.8	210	1 HIS2_VIBCH	Q9ksw7 vibrio chol

c 92 7 0.8 210 1 HIS7 MYCLE
 c 93 7 0.8 211 1 ISPD CHLPN
 c 94 7 0.8 211 1 Y290_WETJA
 c 95 7 0.8 217 1 EVGI_HUMAN
 c 96 7 0.8 219 1 ATP6_ONCMA
 c 97 7 0.8 223 1 ATP6_ONCMY
 c 98 7 0.8 226 1 ATP6_XENLA
 c 99 7 0.8 227 1 ATP6_SALSA
 c 100 7 0.8 228 1 NEUA_HAEIN
 Q9x7b9 mycobacteri
 Q9z7x5 chlamydia p
 Q57738 methanococ
 Q9bze7 homo sapien
 Q36964 oncorhynchu
 P48178 oncorhynchu
 P00849 xenopus lae
 Q35920 salmo salar
 Q57140 haemophilus

ALIGNMENTS

RESULT 1
 SM3A_HUMAN STANDARD; PRT; 771 AA.
 AC Q14563;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 3A precursor (Semaphorin III) (Sema III).
 GN SEMA3A.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=94094332; PubMed=8269517;
 RA Kolodkin A.L., Matthes D.J., Goodman C.S.;
 RT "The semaphorin genes encode a family of transmembrane and secreted
 RT growth cone guidance molecules.";
 RL Cell 75:1389-1399(1993).
 RN [2]
 RP SEQUENCE OF 1-37 FROM N.A.
 RA Woessner J., Minx P., Hinds K., Strommatt C.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 39-182 FROM N.A.
 RA Rohlfing T., Tin-Wollam A.M., Duckels G.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
 CC CONES. COULD SERVE AS A LIGAND THAT GUIDES SPECIFIC GROWTH CONES
 CC BY A MOTILITY-INHIBITING MECHANISM. BINDS TO THE COMPLEX
 CC NEUROPILIN-1/PLEXIN-1 (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
 CC THIRD OF THE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; L26081; AAA65938.1; -;
 DR EMBL; AC004451; -; NOT_ANNOTATED_CDS.
 DR EMBL; AC004848; AAC78622.1; -;
 DR PIR; D49423; D49423.
 DR Genew; HGNC:10723; SEMA3A.
 DR MIM; 603961; -;
 DR GO; GO:0005576; C:extracellular; TAS.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003006; Ig_MHC.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.

DR Pfam; PF00047; ig; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 20
 FT CHAIN 21 771
 FT DOMAIN 240 538
 FT DOMAIN 580 664
 FT DOMAIN 727 769
 FT DISULFD 649 722
 FT CARBOHYD 53 53
 FT CARBOHYD 125 125
 FT CARBOHYD 590 590
 SQ SEQUENCE 771 AA; 88889 MW; 9985F8D3EAED8456 CRC64;
 Alignment Scores:
 Pred. No.: 0 Length: 771
 Score: 771.00 Matches: 771
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 86.15% Indels: 0
 DB: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x SM3A_HUMAN (1-771)
 QY 200 ATGGGCTGGTTAACTAGGATTGTCTGCTTTCTGGGAGTATTACTACAGCAAGCA 259
 Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThAlaAlaAla 20
 QY 260 AACTATCAGATGGGAAGAACAAATGTGCCAGGCTGAAATTTATCTCTACAAAGAATGTG 319
 Db 21 AsnTyrGlnAsnGlyLysAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
 QY 320 GAATCCCAACATGTGATCATTCTCAATGGCTTGGCCACAGCTCCAGTTATCATCCTTC 379
 Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
 QY 380 CTTTGGATGAGGAACGGAGTAGGCTGTATCTTGGAGCAAGGATCATCATTTTCATTC 439
 Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
 QY 440 GACCTGTTATATATCAGGATTTTCAAAAGATTGTGGCCAGTATCTTACACCAAGA 499
 Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArg 100
 QY 500 GATGAATGCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTCTAAATTTTCATCAAGTA 559
 Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPhelLysVal 120
 QY 560 CTTAAGGCATATAATACAGACTCATTGTGACGCTGTGGAACGGGGCTTTTCATCCAATT 619
 Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
 QY 620 TGCACCTACATTAATAATGGACATCATCTCCTGAGGACAATATTTTAACTGGAGAACTCA 679
 Db 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
 QY 680 CATTTTGAACACGGCTGGGAAGAGTCCATATGACCTCCCTTAAGCTGTGCAGAGCATCCCTT 739
 Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThAlaSerLeu 180
 QY 740 TTAATAGATGGAGAAATATATCTCTGGAACCTGCAGCTGATTTTATGGGCGAGACTTTCCT 799
 Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
 QY 800 ATCTTCGAACTCTTGGCCACCACCCCAATCAGGACAGACAGCATGATTCAGGTTGG 859
 Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220

QY 850 CTCATGATCCAAAGTTTCATTAGTCCGCCACCTCATCTCAGAGAGTGACAATCTCGAAGAT 919
 Db 221 LeuAsnAspProLysPhePheIleSerAlaHisLeuSerGluSerAspAsnProGluAsp 240
 QY 920 GACAAAGTATCTTTTCTCCGTCGAAATGCAATAGATGAGACACTCTCGGAAAAGCT 979
 Db 241 AspylsvaltyrPhePhePheargGluAsnAlaIleAspGlyGluHiserGlylysAla 260
 QY 980 ACTCAGCGTAGAATAGTGCAGATATGCAAGATGACTTTGGAGGCGCACAGAAGTCTCGTG 1039
 Db 261 ThrHisAlaArgIleGlyGlnleCyslysAsnAspPheGlyGlyHisArgSerLeuVal 280
 QY 1040 AATAAATGGACAACTTCTCCTCAAGCTCGTCTGATTTCTCAGTCCGCCAGGTCCTCAAGTGC 1099
 Db 281 AsnLysTrpThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
 QY 1100 ATTGACACTCATTTTGTGATGACTGCAGATGATGATTCCTAATGCACTTTAAAGATCCTAAA 1159
 Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPhelyAspProlys 320
 QY 1160 AATCCAGTTGTATATGAGTGTGTTACGACTTCCAGTAACTTTTCAAGGGATCAGCGGTG 1219
 Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
 QY 1220 TGTATGTATAGCATGAGTGTGAGAGGGTGTCTTCTGGTCCATATGCCACAGGGAT 1279
 Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
 QY 1280 SGACCCAACTATCAATGGTGGCTTATCAAGAGAGAGTCCCTTATCCACGCCCGAGACT 1339
 Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
 QY 1340 TGTCTCCAGCAAACTTTGGTGGTTTTCAGTCTCAAAAGGACCTTCTCGATGATGTTATA 1399
 Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
 QY 1400 ACCTTTGCAAGAGTCATCCAGCCCATGATCAATCAGTGTTCCTATCAACAATCGCCCA 1459
 Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
 QY 1460 ATAGTGATCAAAACCGATTAATATCAATTTACAAATTTGCTAGACCGAGTGGAT 1519
 Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
 QY 1520 GCAGAAAGTACACAGTATGATGTTATGTTTTCGGAACAGATGTTGGACCGTCTCTTAAA 1579
 Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
 QY 1580 GTAGTTTCAATTCCTAAGGAGACTTTGGTATGATTTAGAGAGGTTCTCTCGAAGAAATG 1639
 Db 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluValLeuLeuGluMet 480
 QY 1640 ACAGTTTTCGGAAACCGACTGCTATTTTCAGCAATGGAGCTTCCACTAAGCAGCAACAA 1699
 Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
 QY 1700 CTATATATTTGGTTCAACCGCTGGGTTGCTCCAGCTCCCTTTACACCGGTGTGATTTTAC 1759
 Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
 QY 1760 GGAAGAGCGTGTGCTGAGTGTGCTCCCGAGACCTTACTGTGCTGGATGGTGTCT 1819
 Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
 QY 1820 GCATGTTCTCGTATTTTCCACTGCAAGAGAGCGCACAGACGACAGATATAGAAT 1879
 Db 541 AlaCysSerArgTyrPheProThrAlaLysArgThrArgGlnAspIleArgAsn 560
 QY 1880 GGAGACCCACTGACTCTGTTTACAGCTTACACCATGATATACCATGGCCACAGCCCT 1939
 Db 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580
 QY 1940 GAAGAGAGAATCATCTATGTTGTAGAGAAATAGTAGCACATTTTGGAAATGCAGTCCGAAG 1999

Db 581 GluGluArgIleIleTyrGlyValGluAsnSerThrPheLeuGluCysSerProlys 600
 QY 2000 TCGCAGAGAGCGCTGCTTATTTGGCAATTCAGAGGCGCAATGAAGAGCGCAAGAGAGAG 2059
 Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgGlnGluArgGlyGluGlu 620
 QY 2060 ATCAGAGTGGATGATCATATCATCAGGACAGATCAAGGCGCTTCTGCTAGCTAGTCTCAA 2119
 Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuLeuArgSerLeuGln 640
 QY 2120 CAGAAGATTTCAGCAATTTACTCTGCGATCGCGTGGAAACATGGGTTTCATACAACTCTT 2179
 Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
 QY 2180 CTTAAGGTAAACCTCGGAAGTCAATGACACAGAGCAATTTGGAAGAACTTCTTCATAAAGAT 2239
 Db 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLysAsp 680
 QY 2240 GATCATCGAGATGCTCTAAGACCAAGAAATGCTCAATAGCATGACACCTAGCCAGAG 2299
 Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
 QY 2300 GTCTGGTACAGAGACTTCATGCGACTCATCAACACCCCAATCTCAACACATGATGATGAG 2359
 Db 701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
 QY 2360 TTCTGTCAACAAGTTTCGAAAAAGGACCGAAAAACAACGTCGGCAAGGCCAGGACATACC 2419
 Db 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
 QY 2420 CCAGGGAACAGTAACTAAGAGAGTCAAGTCAAGAAATTAAGAAAGGTAGAAACAGGAGG 2479
 Db 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
 QY 2480 ACCACAGAAATTTAGAGGGGACCCAGGAGTGTC 2512
 Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 2
 SM3A_MOUSE
 ID SM3A_MOUSE STANDARD; PRT; 772 AA.
 AC 008655; Q62180; Q62215;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 3A precursor (Semaphorin III) (Sema III) (Semaphorin D) (Sema D).
 DE SEMA3A OR SEMAD OR SEMD.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NMRI; TISSUE=Embryo;
 RX MEDLINE=95267431; PubMed=7748561;
 RA Puschel A.W., Adams R.H., Betz H.;
 RT "Murine semaphorin D/collapsin is a member of a diverse gene family
 and creates domains inhibitory for axonal extension.";
 RL Neuron 14:941-948(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97470885; PubMed=9331345;
 RA Taniguchi M., Yuasa S., Fujisawa H., Naruse I., Saga S., Mishina M.,
 RA Yagi T.;
 RT "Disruption of semaphorin III/D gene causes severe abnormality in
 peripheral nerve projection.";
 RL Neuron 19:519-530(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Kimura T., Fishman M.C.;
 RT "cDNA sequence of mouse collapsin/semaphorin III.";

Submitted (FEB-2002) to the EMBL/GenBank/DDBJ databases.
 [4]
 SEQUENCE OF 107-772 FROM N.A.
 TISSUE=Petal brain;
 MEDLINE=95267432; PubMed=7748562;
 RA Messersmith E.K., Leonardo E.D., Shatz C.J., Tessier-Lavigne M.,
 Goodman C.S., Kolodkin A.L.;
 RA "Semaphorin III can function as a selective chemorepellent to pattern
 RT sensory projections in the spinal cord";
 RL Neuron 14:949-959(1995).
 CC -!- FUNCTION: PLAY A ROLE IN GROWTH CONES GUIDANCE. MAY FUNCTION TO
 CC PATTERN SENSORY PROJECTIONS BY SELECTIVELY REPELLING AXONS THAT
 CC NORMALLY TERMINATE DORSALLY.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- DEVELOPMENTAL LOCATION: EXPRESSED EARLY IN EMBRYONIC DEVELOPMENT
 CC (ELL) IN DISTINCT REGIONS OF THE NEUROECTODERM AND MESODERM.
 CC EXPRESSION BECAME MORE EXTENSIVE AT LATER STAGES.
 CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
 CC THIRD OF THE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X85993; CAA59985.1; -
 DR EMBL; D85028; BAA19773.1; -
 DR EMBL; L41541; AAL77611.1; -
 DR EMBL; L40484; AAA73934.1; -
 DR PIR; I48747; I48747.
 DR PIR; I58169; I58169.
 DR MGD; MGI:107558; Sema3a.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS50835; IG LIKE; 1.
 DR Signal; Immunglobulin domain; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 20 POTENTIAL.
 FT CHAIN 21 772 SEMAPHORIN 3A.
 FT DOMAIN 240 538 SEMA.
 FT DOMAIN 579 665 IG-LIKE C2-TYPE.
 FT DOMAIN 728 770 ARG/LYS-RICH (BASIC).
 FT DISULFID 650 723 BY SIMILARITY.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 125 125 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 591 591 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 193 193 D -> N (IN REF. 4).
 FT CONFLICT 207 207 H -> D (IN REF. 1).
 FT CONFLICT 253 253 D -> G (IN REF. 1).
 FT CONFLICT 352 352 F -> L (IN REF. 4).
 FT CONFLICT 403 403 A -> G (IN REF. 1).
 FT CONFLICT 571 572 QH -> ED (IN REF. 1).
 FT CONFLICT 616 620 EDKRE -> RRSKR (IN REF. 1).
 FT CONFLICT 623 623 R -> K (IN REF. 4).
 SQ SEQUENCE 772 AA; 88799 MW; E89A08528B10AEC3 CRC64;

Alignment Scores:

Pred. No.: 2,046-89

Length: 772

Matches: 92

Score: 92.00

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 10.28% Indels: 0
 Dbs: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x SM3A_MOUSE (1-772)
 QY 1169 GTATATGAGTGGTTTACGACTTCCAGTAACATTTTCAAGGGATCAGCGGTGTGTATGTAT 1228
 Db |||||||
 QY 324 VallyrGlyValPheThrSerSerAsnlePheLysGlySerAlaValCysMetLyr 343
 Db |||||||
 QY 1229 AGCATGAGTGTGAGAGGGTGTCTCTTGGTCCATATGCCACAGGATGGACCCCAAC 1288
 Db |||||||
 Db 344 SerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAspGlyProAsn 363
 QY 1289 TATCAATGGTGGCTTATCAAGAGAGTCCCTTATCCAGCGCAGGACTTGTCCCGAC 1348
 Db |||||||
 Db 364 TyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThrCysProSer 383
 QY 1349 AAAACATTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1408
 Db |||||||
 Db 384 LysThrPheGlyGlyPheAspSerThrLysAspLeuProAspValIleThrPheAla 403
 QY 1409 AGAAGTCATCCAGCCATGTACATCCAGTGTCTTCT 1444
 Db |||||||
 Db 404 ArgSerHisProAlaMetTyrAsnProValPhePro 415
 RESULT 3
 SM3A RAT
 ID SM3A RAT STANDARD; PRT; 772 AA.
 AC Q63548;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 3A precursor (Semaphorin III) (Sema III).
 GN SEMA3A.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Brain;
 RX MEDLINE=97073089; PubMed=8915837;
 RA Giger R.J., Wolfer D.P., De Wit G.M.J., Verhaagen J.;
 RT "Anatomy of rat semaphorin III/collapsin-1 mRNA expression and
 RL relationship to developing nerve tracts during neuroembryogenesis.";
 RL J. Comp. Neurol. 375:378-392(1996).
 CC -!- FUNCTION: MAY BE INVOLVED IN GUIDING GROWING AXONS TOWARDS THEIR
 CC TARGETS BY FORMING A MOLECULAR BOUNDARY THAT INSTRUCTS AXONS TO
 CC ENGAGE IN THE FORMATION OF SPECIFIC NERVE TRACTS. BINDS TO
 CC NEUROPILIN.
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -!- DEVELOPMENTAL STAGE: AT E-11, EXPRESSION WAS RESTRICTED TO THE
 CC OLFACTORY PIT, THE BASAL AND ROSTRAL SURFACE OF THE TELENCEPHALIC
 CC VESICLE, THE EYE ANLAGE, THE EPITHELIUM OF RATHKE'S POUCH, AND
 CC SOMITES. AT LATER DEVELOPMENTAL STAGES, IT WAS WIDELY DISTRIBUTED
 CC IN NEURONAL AS WELL AS IN MESENCHYMAL AND EPITHELIAL STRUCTURES
 CC OUTSIDE THE NERVOUS SYSTEM. AFTER BIRTH, MESENCHYMAL LEVELS
 CC DECREASED RAPIDLY AND EXPRESSION BECAME RESTRICTED TO SPECIFIC
 CC SETS OF NEURONS IN THE CNS. IN THE MATURE CNS, IT IS DETECTABLE IN
 CC MITRAL CELLS, NEURONS OF THE ACCESSORY BULB AND CEREBRAL CORTEX,
 CC CEREBELLAR PURKINJE CELLS, AS WELL AS A SUBSET OF CRANIAL AND
 CC SPINAL MOTONEURONS.
 CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
 CC THIRD OF THE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; X95286; CAA64607.1; --
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig.MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR001627; Sema.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; Ig; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS00835; IG LIKE; 1.
KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 20
FT CHAIN 1 20
FT DOMAIN 21 772
FT DOMAIN 240 538
FT DOMAIN 577 665
FT DOMAIN 728 770
FT DISULFID 650 723
FT CARBOHYD 53 53
FT CARBOHYD 125 125
FT CARBOHYD 591 591
SQ SEQUENCE 772 AA; 88808 MW; 240907812FF9F2D2 CRC64;

Alignment Scores:
Pred. No.: 2.04e-89 Length: 772
Score: 92.00 Matches: 92
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.28% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3A_RAT (1-772)

QY 1169 GTATGAGTGTGTTTACGACTTCAGTAACATTTTCAAGGATCAGCGTGTATGAT 1228
DB 324 ValTyrGlyValPheThrSerSerAsnIlePheGlySerAlaValCysMetTyr 343
QY 1229 AGCATGAGTGTGAGAGGGTGTCTTGGTCCATATGCCAGGATGACCCCAAC 1288
DB 344 SerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAspGlyProAsn 363
QY 1289 TATCATGGGTGCTTATCAAGAGAGTCCCTATCCAGCGGAGCACTTGTCCAGC 1348
DB 364 TyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThrCysProSer 383
QY 1349 AAAACATTTGGTGTGTTGACTTACAAAGACCTTCTGTATGATGTTTATACCTTTGCA 1408
DB 384 LysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIleThrPheAla 403
QY 1409 AGAAGTCATCCAGGATGATCAATCCAGTGTCTCT 1444
DB 404 ArgSerHisProAlaMetTyrAsnProValPhePro 415

RESULT 4

SM3A_CHICK
ID SM3A_CHICK STANDARD; PRT; 772 AA.
AC Q90607;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Senaphorin 3A precursor (Collapsin-1) (COLL-1).
OS SEMA3A OR COLL1.
GN Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.

NCBI_TaxID=9031;
[1]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 322-329; 362-372; 395-403 AND
RP 666-680.
RC TISSUE=Brain;
RX MEDLINE=94006554; PubMed=8402908;
RA Luo Y., Raible D., Raper J.A.;
RT "Collapsin: a protein in brain that induces the collapse and paralysis
of neuronal growth cones.";
RL Cell 75:217-227(1993).
CC -!- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
CONES. COULD SERVE AS A LIGAND THAT GUIDES SPECIFIC GROWTH CONES
BY A MOTILITY-INHIBITING MECHANISM. BINDS TO NEUROFILIN.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- TISSUE SPECIFICITY: EXPRESSED AT RELATIVELY HIGH LEVELS IN BRAIN
AND MUSCLE, MODERATE LEVELS IN LUNG, BURSA, AND HEART AND
CC VIRTUALLY ABSENT IN LIVER. COLLAPSN-1, -2, -3, AND -5 BIND TO
CC OVERLAPPING BUT DISTINCT AXON TRACTS.
CC -!- DOMAIN: STRONG BINDING TO NEUROFILIN IS MEDIATED BY THE CARBOXY
THIRD OF THE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U02528; AAC59638.1; --
DR PIR; A49069; A49069.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig.MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR001627; Sema.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; Ig; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS00835; IG LIKE; 1.
KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 22
FT CHAIN 1 22
FT DOMAIN 23 772
FT DOMAIN 240 538
FT DOMAIN 576 665
FT DOMAIN 728 770
FT DISULFID 650 723
FT CARBOHYD 53 53
FT CARBOHYD 125 125
FT CARBOHYD 591 591
SQ SEQUENCE 772 AA; 88867 MW; E91E09DE0CC940AC CRC64;

Alignment Scores:

Pred. No.: 5.33e-53 Length: 772
Score: 58.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.48% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3A_CHICK (1-772)

QY 971 GGAAGCTACTCAGCTAGATAGTGCATATGCAAGATGACCTTTGGAGGCACAGA 1030
DB 258 GlyylsalathrHisalaArgIleGlyIleCysLysAsnAspPheGlyGlyHisArg 277
QY 1031 AGTCTGTGTAATAATGGACAACATCTCTCAAAGCTCGTCTGATTGCTCAGTGCCAGGT 1090

Db 278 SerLeuValAsnLysTrpThrPheLeuLysAlaArgLeuLeuCysSerValProGly 297
 QY 1091 CCAATGGCATTCACACTCATTTTGATGAAGTTCAGTATTCCTATGAAC 1144
 Db 298 ProAsnGlyLeuAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsn 315

RESULT 5

ID SZ1A_BRARE STANDARD; PRT; 860 AA.
 AC Q9W7J1;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin Z1A precursor (Semaphorin 1A) (Sema-Z1A).
 GN SEMAZ1A OR SEMAZ1A.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=99313409; PubMed=10386838;
 RA Yee C.S., Chandrasekhar A., Halloran M.C., Shoji W., Warren J.T.,
 RA Kuwada J.Y.;
 RT "Molecular cloning, expression, and activity of zebrafish semaphorin
 Z1a.";
 RL Brain Res. Bull. 48:581-593(1999).
 CC -!- FUNCTION: MAY INFLUENCE OUTGROWTH BY A VARIETY OF GROWTH CONES
 CC INCLUDING THOSE OF THE POSTERIOR LATERAL LINE GANGLION.
 CC -!- SUBCELLULAR LOCATION: Secreted (by similarity).
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN HIGHLY SPECIFIC PATTERNS WITHIN
 CC THE DEVELOPING EMBRYO.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF086761; AAD43964.1; -;
 CC ZFIN; ZDB-GENE-991209-3; sema3aa.
 CC InterPro; IPR007110; IG-like.
 CC InterPro; IPR003599; IG.
 CC InterPro; IPR003006; IG_MHC.
 CC InterPro; IPR003659; Plexin-like.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF00047; IG; 1.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00409; IG; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC SMART; PS50835; IG LIKE; 1.
 CC Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 CC Developmental protein; Glycoprotein.
 CC SIGNAL 1 17
 CC CHAIN 18 860
 CC DOMAIN 241 539
 CC DOMAIN 579 668
 CC DOMAIN 722 858
 CC DISULFID 652 717
 CC CARBOHYD 53 53
 CC CARBOHYD 126 126
 CC CARBOHYD 593 593
 CC SEQUENCE 860 AA; 97263 MW; 5FD4C12194F5165C CRC64;

Alignment Scores:
 Pred. No.: 2, 28e-28 Length: 860
 Score: 35.00 Matches: 35
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.91% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SZ1A_BRARE (1-860)

QY 1307 CAAGAGAGATCCCTATCCACGGCCAGAACTGTCCAGCAAAACATTGGTGTGTTT 1366
 Db 371 GlnGlyArgValProTyrProArgProGlyThrCysProSerLysThrPheGlyGlyPhe 390
 QY 1367 GACTCTACAAAGGACCTTCCTGATGATGTTTATACCTTTGCAAGA 1411
 Db 391 AspSerThrLysAspLeuProAspValIleThrPheAlaArg 405

RESULT 6

ID SZ1B_BRARE STANDARD; PRT; 778 AA.
 AC Q9W686;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin Z1B precursor (Semaphorin 1B) (Sema-Z1B).
 GN SEMAZ1B OR SEMAZ1B.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99425174; PubMed=10495275;
 RA Roos M., Schachner M., Bernhardt R.R.;
 RT "Zebrafish semaphorin Z1b inhibits growing motor axons in vivo.";
 RL Mech. Dev. 87:103-117(1999).
 CC -!- FUNCTION: MIGHT NORMALLY INFLUENCE THE MIDSEGMENTAL PATHWAY CHOICE
 CC OF THE VENTRALLY EXTENDING MOTOR AXONS BY CONTRIBUTING TO A
 CC REPULSIVE DOMAIN IN THE POSTERIOR SOMITE.
 CC -!- SUBCELLULAR LOCATION: Secreted (by similarity).
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN RHOMBOMERES THREE AND FIVE, AND
 CC IN THE POSTERIOR HALF OF NEWLY FORMED SOMITES WHICH IS AVOIDED BY
 CC VENTRALLY EXTENDING MOTOR AXONS.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF083382; AAD28103.1; -;
 CC ZFIN; ZDB-GENE-991209-6; sema3ab.
 CC InterPro; IPR007110; IG-like.
 CC InterPro; IPR003599; IG.
 CC InterPro; IPR003006; IG_MHC.
 CC InterPro; IPR003659; Plexin-like.
 CC InterPro; IPR002165; Plexin-repeat.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF00047; IG; 1.
 CC Pfam; PF01437; PSI; 1.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00409; IG; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC SMART; PS50835; IG LIKE; 1.
 CC Signal; Immunoglobulin domain; Multigene family; Neurogenesis;

SQ

KW Developmental protein; Glycoprotein.
FT SIGNAL; 1 17 POTENTIAL.
FT CHAIN 18 778 SEMAPHORIN 21B.
FT DOMAIN 241 539 SEMA.
FT DOMAIN 579 668 IG-LIKE C2-TYPE.
FT DOMAIN 721 776 ARG/LYS-RICH (BASIC).
FT DISULFID 652 716 BY SIMILARITY.
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 127 127 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 593 593 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 778 AA; 88904 MW; 4D36F432A21895 CRC64;

Alignment Scores:
Pred. No.: 1.39e-16 Length: 778
Score: 24.00 Matches: 24
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.68% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x S21B_BRARE (1-778)
QY 1823 TGTTCTCGCTATTTCCCACTGCAAGAGACGACACAGACAGATATTAAGAAATGGA 1882
Db 543 CysSerArgTyrPheProThrAlaLysArgThrArgGlnAspIleArgAsnGly 562

QY 1883 GACCACTGACT 1894
Db 563 AspProLeuThr 566

RESULT 7
SM3B_MOUSE
ID SM3B_MOUSE STANDARD; PRT; 748 AA.
AC Q62177;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3B precursor (Semaphorin A) (Sema A).
GN SEMA3B OR SEMAA OR SEMA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NMRI; TISSUE=Embryo;
RX MEDLINE=95267431; PubMed=7748561;
RA Puechel A.W., Adams R.H., Betz H.;
RT "Murine semaphorin D/collapsin is a member of a diverse gene family
and creates domains inhibitory for axonal extension.";
RL Neuron 14:941-948(1995).
CC -!- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE EMBRYO. LOW
LEVELS FOUND BETWEEN DAYS 10-12. EXPRESSION PEAKS ON DAY 13 WITH
MODERATE LEVELS FROM THEN UNTIL BIRTH.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; X85990; CAA59982.1; -.
DR PIR; I48744; I48744.
DR MGI; MGI:107561; Sema3b.
DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR001627; Sema.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS00835; IG LIKE; 1.
KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
Developmental protein; Glycoprotein.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 748 SEMAPHORIN 3B.
FT DOMAIN 55 496 SEMA.
FT DOMAIN 561 659 IG-LIKE C2-TYPE.
FT DOMAIN 698 702 POLY-GLY.
FT DOMAIN 723 743 ARG-RICH (BASIC).
FT DISULFID 643 709 BY SIMILARITY.
FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 748 AA; 82894 MW; 1866B7D2397C9305 CRC64;

Alignment Scores:
Pred. No.: 2.68e-12 Length: 748
Score: 20.00 Matches: 20
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.23% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3B_MOUSE (1-748)
QY 1766 GCGTGTGCTGAGTGTGCTCGCCGAGACCTTACTGTGTGGATGTTCTGCATGT 1825
Db 521 AlaCysAlaGluCysCysLeuAlaArgAspProIrrCysAlaIrrPaspGlySerAlaCys 540

RESULT 8
SM22_BRARE
ID SM22_BRARE STANDARD; PRT; 764 AA.
AC Q9W6G6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Semaphorin Z2 precursor (Semaphorin 2) (Sema-22).
GN SEMA22 OR SEMA2.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99112778; PubMed=9915572;
RA Halloran M.C., Severance S.M., Yee C.S., Gemza D.L., Raper J.A.,
RA Kuwada J.Y.;
RT "Analysis of a Zebrafish semaphorin reveals potential functions in
vivo.";
RL Dev. Dyn. 214:13-25(1999).
CC -!- FUNCTION: MAY PLAY A ROLE IN THE GUIDANCE OF SEVERAL AXON
PATHWAYS.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- DEVELOPMENTAL STAGE: EXPRESSED IN A DYNAMIC AND RESTRICTED PATTERN
DURING THE PERIOD OF AXON OUTGROWTH.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

CC EMBL; X85990; CAA59982.1; -.
DR PIR; I48744; I48744.
DR MGI; MGI:107561; Sema3b.
DR InterPro; IPR007110; Ig-like.

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

CC -----
CC EMBL; AF124485; AAD21310.1; -.
CC ZFIN; ZDB-GENE-990715-2; sema2.
CC InterPro; IPR003599; Ig.
CC InterPro; IPR003659; Plexin-like.
CC InterPro; IPR001627; Sema.
CC Pfam; PF01403; Sema; 1.
CC SMART; SM00409; Ig; 1.
CC SMART; SM00423; PSI; 1.
CC SMART; SM00630; Sema; 1.
CC Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 41
FT CHAIN 42 764 SEMAPHORIN 22.
FT DOMAIN 261 559 SEMA.
FT DOMAIN 661 740 IG-LIKE C2-TYPE.
FT DOMAIN 741 762 ARG/LYS-RICH (BASIC).
FT DISULFID 668 733 BY SIMILARITY.
FT CARBOHYD 143 143 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 490 490 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 610 610 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 764 AA; 87859 MW; A3ED95C2C479D7AE CRC64;

Alignment Scores:
Pred. No.: 5,16e-08 Length: 764
Score: 16.00 Matches: 16
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.79% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SMZ2_BARE (1-764)

QY 1769 TGTGCTGAGTGTGCTGCCCGACACCTTACTGCTGGATGGT 1816
Db 545 CysAlaGluCysCysLeuAlaArgAspProTyCysAlaTrpAspGly 560

RESULT 9
SM3B_HUMAN
ID SM3B_HUMAN STANDARD; PRT; 749 AA.
AC Q13214; Q8TDV7; Q93018;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3B precursor (Semaphorin V) (Sema V) (Sema A(V)).
GN SEMA3B OR SEMA5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96210603; PubMed=8633026;
RA Sekido Y., Bader S., Latif F., Chen J.-Y., Duh F.-M., Wei M.-H.,
RA Albanesi J.P., Lee C.-C., Lerman M.I., Minna J.D.;
RT "Human semaphorins A(V) and IV reside in the 3p21.3 small cell lung
RT cancer deletion region and demonstrate distinct expression patterns.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:4120-4125(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Dante M., Wamsley P.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Koyama N.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP VARIANTS NSCLC CVS-348; HIS-397 AND ILE-415.
RX MEDLINE=20535986; PubMed=11085536;
RA Lerman M.I.; Minna J.D.;

```

```

RT "The 630-kb lung cancer homozygous deletion region on human chromosome
RT 3p21.3: identification and evaluation of the resident candidate tumor
RT suppressor genes."
RL Cancer Res. 60:6116-6133(2000).
CC -!- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
CC SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: SECRETED (BY SIMILARITY). ACCUMULATES IN THE
CC ENDOPLASMIC RETICULUM.
CC -!- TISSUE SPECIFICITY: Expressed abundantly but differentially in a
CC variety of neural and nonneural tissues.
CC -!- DISEASE: Defects in SEMA3B are found in non-small cell lung cancer
CC (NSCLC) cell lines.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

Alignment Scores:
Pred. No.: 6,08e-07 Length: 749
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0

DB: 1 0 Gaps: 0

US-09-774-490-1 (1-2709) x SM3B_HUMAN (1-749)

QY 1028 AGAAGTCTGGTGAATAAATGGAACAATTCCTCAAGCTCGTCTG 1072
|||||

Db 277 ArgSerLeuValAsnLysTrpThrThrPheLeuLysAlaArgLeu 291
|||||

RESULT 10

SM3C_CHICK

ID SM3C_CHICK STANDARD; PRT; 751 AA.

AC 042336; Q90664;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Semaphorin 3C precursor (Collapsin-3) (COLL-3).

GN SEMA3C OR COLL3.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;

OC Gallus.

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Fetal brain;

RX MEDLINE=97470887; PubMed=9331347;

RA Luo Y., Shepherd I., Li J., Renzi M.J., Chang S., Raper J.A.;

RT "A family of molecules related to collapsin in the embryonic chick nervous system";

RL Neuron 19:539-545(1997).

RN [2]

RP SEQUENCE OF 237-530 FROM N.A.

RX MEDLINE=95329269; PubMed=7605628;

RA Luo Y., Shepherd I., Li J., Renzi M.J., Chang S., Raper J.A.;

RT "A family of molecules related to collapsin in the embryonic chick nervous system";

RL Neuron 14:1131-1140(1995).

CC -!- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH CONES. COULD POTENTIALLY ACT AS REPULSIVE CUES TOWARD SPECIFIC NEURONAL POPULATIONS. BINDS TO NEUROPILIN.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: COLLAPLIN-1, -2, -3, AND -5 BIND TO OVERLAPPING BUT DISTINCT AXON TRACTS.

CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY THIRD OF THE PROTEIN.

CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.

CC -!- SIMILARITY: Contains 1 Sema domain.

CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AF022946; AAB80951.1; -

CC EMBL; U28241; AAB86897.1; -

CC InterPro; IPR007110; IG-like.

CC InterPro; IPR003599; IG.

CC InterPro; IPR003659; Plexin-like.

CC InterPro; IPR001627; Sema.

CC Pfam; PF01403; Sema; 1.

CC SMART; SM00409; IG; 1.

CC SMART; SM00423; FS1; 1.

CC SMART; SM00630; Sema; 1.

CC PROSITE; PS50835; IG LIKE; 1.

KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis; Developmental protein; Glycoprotein.

FT SIGNAL 1 21 POTENTIAL.

FT CHAIN 22 751 SEMAPHORIN 3C.

FT DOMAIN 237 535 SEMA.

FT DOMAIN 571 655 IG-LIKE C2-TYPE.

FT DOMAIN 718 749 ARG/LYS-RICH (BASIC).

FT DISULFID 643 709 BY SIMILARITY.

FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 123 123 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 268 268 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 465 465 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 585 585 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 586 586 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CONFLICT 238 238 A -> D (IN REF. 2).

FT CONFLICT 359 359 P -> S (IN REF. 2).

FT CONFLICT 404 404 H -> D (IN REF. 2).

SQ SEQUENCE 751 AA; 85433 MW; DB686687D21BD5D8 CRC64;

Alignment Scores:

Pred. No.: 6.08e-07 Length: 751

Score: 15.00 Matches: 15

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.68% Indels: 0

DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3C_CHICK (1-751)

QY 1028 AGAAGTCTGGTGAATAAATGGAACAATTCCTCAAGCTCGTCTG 1072
|||||

Db 274 ArgSerLeuValAsnLysTrpThrThrPheLeuLysAlaArgLeu 288
|||||

RESULT 11

SM3C_HUMAN

ID SM3C_HUMAN STANDARD; PRT; 751 AA.

AC Q99985;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Semaphorin 3C precursor (Semaphorin E) (Sema E).

GN SEMA3C OR SEMAE.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98070817; PubMed=9405678;

RA Yamada T., Endo R., Gotoh M., Hirohashi S.;

RT "Identification of semaphorin E as a non-MDR drug resistance gene of human cancers";

RL Proc. Natl. Acad. Sci. U.S.A. 94:14713-14718(1997).

CC -!- FUNCTION: MAY BE INVOLVED IN DIVERSE CELL SURVIVAL MECHANISMS.

CC -!- SUBCELLULAR LOCATION: Secreted.

CC -!- TISSUE SPECIFICITY: EXPRESSED INTENSELY IN THE HEART, SKELETAL MUSCLE, COLON, SMALL INTESTINE, OVARY, TESTIS, AND PROSTATE. PAINT EXPRESSION UBQUITOUSLY AMONG OTHER ORGANS, INCLUDING BRAIN.

CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.

CC -!- SIMILARITY: Contains 1 Sema domain.

CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AB000220; BAA32398.1; -

CC EMBL; HGNC:10725; SEMA3C.

CC MIM; 602645; -

CC GO; GO:0008151; P:cell growth and/or maintenance; TAS.

CC GO; GO:0009315; P:drug resistance; TAS.

CC GO; GO:0006955; P:immune response; TAS.

CC InterPro; IPR007110; IG-like.

CC InterPro; IPR003599; IG.

```
DR InterPro; IPR003006; Ig MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR001627; Sema.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS00835; IG LIKE; 1.
KW Signal; Immunoglobulin domain; Multigene family; Glycoprotein.
FT SIGNAL 1 20
FT CHAIN 21 751
FT DOMAIN 237 535
FT DOMAIN 571 655
FT DOMAIN 724 745
FT DISULFID 643 709
FT CARBOHYD 81 81
FT CARBOHYD 123 123
FT CARBOHYD 252 252
FT CARBOHYD 268 268
FT CARBOHYD 465 465
FT CARBOHYD 585 585
FT CARBOHYD 586 586
SQ SEQUENCE 751 AA; 85207 MW; AA68A21FEF20C033 CRC64;

Alignment Scores:
Pred. No.: 6.08e-07 Length: 751
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3C_HUMAN (1-751)
Qy 1028 AGAAGTCTGGTGAATAATGACACATTCCTCAAGCTGCTG 1072
Db 274 ArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeu 288

RESULT 12
SM3C_MOUSE
ID SM3C_MOUSE STANDARD; PRT; 751 AA.
AC Q62181;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3C precursor (Semaphorin E) (Sema E).
GN SEMA3C OR SEMAE OR SEME.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=NMRI; TISSUE=Embryo;
RX MEDLINE=95267431; PubMed=7748561;
RA Puschel A.W., Adams R.H., Betz H.;
RT "Murine semaphorin D/collapsin is a member of a diverse gene family
and creates domains inhibitory for axonal extension.";
RL Neuron 14:941-948(1995).
CC -!- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE EMBRYO. MAXIMUM
EXPRESSION BETWEEN DAYS 10-12 WITH MODERATE LEVELS FROM DAY 13
UNTIL BIRTH.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X85994; CAA59986.1; -.
DR PIR; I48748; I48748.
DR MGI; MGI:107557; Sema3c.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003659; Plexin-like.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS00835; IG LIKE; 1.
KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
FT SIGNAL 1 20
FT CHAIN 21 751
FT DOMAIN 237 535
FT DOMAIN 571 655
FT DOMAIN 714 717
FT DOMAIN 724 745
FT DISULFID 643 709
FT CARBOHYD 81 81
FT CARBOHYD 123 123
FT CARBOHYD 252 252
FT CARBOHYD 268 268
FT CARBOHYD 465 465
FT CARBOHYD 585 585
FT CARBOHYD 586 586
SQ SEQUENCE 751 AA; 85259 MW; A9B599SE5E7DB45D8 CRC64;

Alignment Scores:
Pred. No.: 6.08e-07 Length: 751
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3C_MOUSE (1-751)
Qy 1028 AGAAGTCTGGTGAATAATGACACATTCCTCAAGCTGCTG 1072
Db 274 ArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeu 288

RESULT 13
SM3D_HUMAN
ID SM3D_HUMAN STANDARD; PRT; 777 AA.
AC O95025;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3D precursor.
GN SEMA3D.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Submittted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RL Submittted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
CONES. COULD POTENTIALLY ACT AS REPULSIVE CUES TOWARD SPECIFIC
NEURONAL POPULATIONS. BINDS TO NEUROPILIN (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
THIRD OF THE PROTEIN.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
```

CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; AC004957; RAC3081.1; ALT_SEQ.
 DR Genes; HGNC:10726; SEMA3D.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003006; Ig MHC.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001827; Sema.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; Ig; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS00835; IG-LIKE; 1.
 KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 39
 FT CHAIN 40 777 SEMAPHORIN 3D.
 FT DOMAIN 257 555 SEMA.
 FT DOMAIN 592 680 IG-LIKE C2-TYPE.
 FT DOMAIN 739 773 ARG/LYS-RICH (BASIC).
 FT DISULFID 665 731 BY SIMILARITY.
 FT CARBOHYD 139 139 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 607 607 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 724 724 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 777 AA; 89651 MW; 3F7B0D7AF50F53BD CRC64;

Alignment Scores:
 Pred. No.: 7,14e-06 Length: 777
 Score: 14.00 Matches: 14
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.56% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3D_HUMAN (1-777)

QY 1040 AATAATGACACATCTCTCAAGCTCGTCTGATTGCTCA 1081
 |||||
 Db 298 AsnLysTrpThrThrPheLeuLysAlaArgLeuLysCysSer 311

RESULT 14
 SM3D_CHICK
 ID SM3D_CHICK STANDARD; PRT; 761 AA.
 AC Q90663;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 3D precursor (Collapsin-2) (COLL-2).
 GN SEMA3D OR COLL2.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95329269; PubMed=7605628;
 RA Luo Y., Shepherd I., Li J., Renzi M.J., Chang S., Raper J.A.;
 RT "A family of molecules related to collapsin in the embryonic chick
 RT nervous system."
 RL Neuron 14:1131-1140(1995).
 CC -!- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH

CC CONES. COULD POTENTIALLY ACT AS REPULSIVE CUES TOWARD SPECIFIC
 CC NEURONAL POPULATIONS. BINDS TO NEUROPILIN.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: DEVELOPING SPINAL CORD AND DEVELOPING VISUAL
 CC SYSTEM. COLLAPLIN-1, -2, -3, AND -5 BIND TO OVERLAPPING BUT
 CC DISTINCT AXON TRACTS.
 CC -!- DOMAIN: STRONG BINDING TO NEUROPILIN IS MEDIATED BY THE CARBOXY
 CC THIRD OF THE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -----

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

EMBL; U28240; AAA86896.1; --
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR003006; Ig MHC.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; Ig; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS00835; IG-LIKE; 1.
 KW Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 24
 FT CHAIN 25 761 SEMAPHORIN 3D.
 FT DOMAIN 245 543 SEMA.
 FT DOMAIN 552 670 IG-LIKE C2-TYPE.
 FT DOMAIN 727 757 ARG/LYS-RICH (BASIC).
 FT DISULFID 653 719 BY SIMILARITY.
 FT CARBOHYD 127 127 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 595 595 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 761 AA; 87300 MW; 3E09AE3DEA53F46B CRC64;

Alignment Scores:
 Pred. No.: 8.43e-05 Length: 761
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3D_CHICK (1-761)

QY 1778 TGTTGCTGCGCCGAGACCTTACTGTGTTGGATGGT 1816
 |||||
 Db 532 CysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 544

RESULT 15
 SM3E_HUMAN
 ID SM3E_HUMAN STANDARD; PRT; 775 AA.
 AC O15041;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 3E precursor.
 GN SEMA3E OR KIAA0331.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;

RX MEDLINE=97349984; PubMed=9205841;
RA Nagase T., Ishikawa K.-I., Nakajima D., Ohira M., Seki N.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VII.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 4:141-150(1997).
RN [2]
RP SEQUENCE OF 1-38 FROM N.A.
RA Holmes A., Johnson D., Elliot G., Burkhardt J.;
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AB002329; BAA20789.1; -;
CC EMBL; AC004954; AAC69513.1; -;
CC Genew; HGNC:10727; SEMA3E
CC GO; GO:0005615; C:extracellular space; TAS.
CC GO; GO:0007048; P:oncogenesis; TAS.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003599; Ig.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR003659; Plexin-like.
CC InterPro; IPR001627; Sema.
CC Pfam; PF00047; Ig; 1.
CC SMART; SM00403; Sema; 1.
CC SMART; SM00423; PSI; 1.
CC SMART; SM00630; Sema; 1.
CC SMART; SM00630; Sema; 1.
CC PROSITE; PS50835; IG_LIKE; 1.
CC Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 775 SEMAPHORIN 3E.
FT DOMAIN 241 540 SEMA.
FT DOMAIN 581 669
FT DOMAIN 737 770 ARG/LYS-RICH (BASIC).
FT DOMAIN 737 770 BY SIMILARITY.
FT DISULFID 654 729
FT CARBOHYD 44 44 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 126 126 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 595 595 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 596 596 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 775 AA; 89227 MW; CD6079C1AB48F779 CRC64;

Alignment Scores:
Pred. No.: 8,41e-05 Length: 775
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3E_HUMAN (1-775)
QY 1778 TTTGCTCCGCCGACAGCCCTTACTGTGCTGGGTGGT 1816
Db 529 CysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 541

RESULT 16
SM3E_MOUSE STANDARD; PRT; 775 AA.
ID SM3E_MOUSE
AC P70275; O09078; O09079;

DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 3E precursor (Semaphorin H) (Sema H).
GN SEMA3E OR SEMAR OR SEMH.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/C;
RX MEDLINE=98175564; PubMed=9515811;
RA Christensen C.R.L., Klingelhofer J., Tarabykina S., Hulgaard E.F.,
RA Kramarov D., Lukanidin E.;
RT "Transcription of a novel mouse semaphorin gene, M-semaH, correlates
RT with the metastatic ability of mouse tumor cell lines.";
RL Cancer Res. 58:1238-1244(1998).
RN [2]
RP REVISIONS
RA Christensen C.R.L.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: MAY BE INVOLVED IN EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: DEVELOPING LUNGS, DEVELOPING SKELETAL
CC ELEMENTS AND VENTRAL HORNS OF THE DEVELOPING NEURAL TUBE.
CC CORRELATES POSITIVELY WITH TUMOR PROGRESSION.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z80941; CAB02590.1; -;
CC EMBL; Z93947; CAB07987.1; ALT SEQ.
CC EMBL; Z93948; CAB07988.1; ALT_SEQ.
CC MGD; MGI:1340034; Sema3e.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003599; Ig.
CC InterPro; IPR003006; Ig_MHC.
CC InterPro; IPR003659; Plexin-like.
CC InterPro; IPR001627; Sema.
CC Pfam; PF00047; Ig; 1.
CC SMART; SM00403; Sema; 1.
CC SMART; SM00423; PSI; 1.
CC SMART; SM00630; Sema; 1.
CC PROSITE; PS50835; IG_LIKE; 1.
CC Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 775 SEMAPHORIN 3E.
FT DOMAIN 241 540 SEMA.
FT DOMAIN 581 669 ARG/LYS-RICH (BASIC).
FT DOMAIN 737 770 BY SIMILARITY.
FT DISULFID 654 729
FT CARBOHYD 44 44 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 126 126 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 175 175 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 596 596 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 775 AA; 89503 MW; BA41690093D28F40 CRC64;

Alignment Scores:
Pred. No.: 8,41e-05 Length: 775
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3E_MOUSE (1-775)
 QY 1778 TGTTCCTCGCCGAGACCTTACTGTCTGGGATGCT 1816
 DB 529 CysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 541

RESULT 17
 SM3E_CHICK
 ID SM3E_CHICK STANDARD; PRT; 785 AA.
 AC 042237; Q90666;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Semaphorin 3E precursor (Collapsin-5) (COLL-5).
 GN SEMA3E OR COLL5
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=97470887; PubMed=9331347;
 RA Feiner L., Koppel A.M., Kobayashi H., Raper J.A.;
 RT "Secreted chick semaphorins bind recombinant neuropilin with similar
 RT affinities but bind different subsets of neurons in situ."
 RL Neuron 19:539-545(1997).
 RN [2]
 RP SEQUENCE OF 244-543 FROM N.A.
 RX MEDLINE=95329269; PubMed=7605628;
 RA Luo Y., Shepherd I., Li J., Renzi M.J., Chang S., Raper J.A.;
 RT "A family of molecules related to collapsin in the embryonic chick
 RT nervous system."
 RL Neuron 14:1131-1140(1995).
 CC -!- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
 CC CONES. COULD POTENTIALLY ACT AS REPULSIVE CUES TOWARD SPECIFIC
 CC NEURONAL POPULATIONS. BINDS TO NEUROFILIN.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: COLLAPSIN-1, -2, -3, AND -5 BIND TO
 CC OVERLAPPING BUT DISTINCT AXON TRACTS.
 CC -!- DOMAIN: STRONG BINDING TO NEUROFILIN IS MEDIATED BY THE CARBOXY
 CC THIRD OF THE PROTEIN.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF022947; AAB80952.1; -;
 DR EMBL; U28243; AAB86899.1; -;
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR Signal; Immunoglobulin domain; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 1 25 SEMAPHORIN 3E.
 FT DOMAIN 26 785 SEMA.
 FT DOMAIN 244 544 SEMA.
 FT DOMAIN 651 740 IG-LIKE C2-TYPE.
 FT DOMAIN 741 780 ARG/LYS-RICH (BASIC).

FT DISULFID 658 733 BY SIMILARITY.
 FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 130 130 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 600 600 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 246 246 N -> D (IN REF. 2).
 FT CONFLICT 248 248 V -> I (IN REF. 2).
 FT CONFLICT 250 250 L -> F (IN REF. 2).
 SQ SEQUENCE 785 AA; 90978 MW; E551EBF717630632 CRC64;

Alignment Scores:
 Pred. No.: 8.41e-05 Length: 785
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3E_CHICK (1-785)
 QY 1778 TGTTCCTCGCCGAGACCTTACTGTCTGGGATGCT 1816
 DB 533 CysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 545

RESULT 18
 SM3F_HUMAN
 ID SM3F_HUMAN STANDARD; PRT; 785 AA.
 AC Q13275; Q13274; Q13372; Q15704;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 3F precursor (Semaphorin IV) (Sema IV) (Sema III/F).
 GN SEMA3F.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal brain;
 RX MEDLINE=96226360; PubMed=8649831;
 RA Roche J., Boldog F., Robinson M., Robinson L., Varella-Garcia M.,
 RA Swanton M., Waggoner B., Fishel R., Franklin W., Gemmill R.,
 RA Drabkin H.;
 RT "Distinct 3p21.3 deletions in lung cancer and identification of a new
 RT human semaphorin."
 RL Oncogene 12:1289-1297(1996).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANT MET-503.
 RC TISSUE=Fetal brain;
 RX MEDLINE=96230324; PubMed=8786119;
 RA Xiang R.-H., Hensel C.H., Garcia D.K., Carlson H.C., Kok K.,
 RA Daly M.C., Korbacher K., van den Berg A., Veldhuis P., Buys C.H.C.M.,
 RA Naylor S.L.;
 RT "Isolation of the human semaphorin III/F gene (SEMA3F) at chromosome
 RT 3p21, a region deleted in lung cancer."
 RL Genomics 32:39-48(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Nelson J., Biewald T.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 394-436 FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=96210603; PubMed=8633026;
 RA Sekido Y., Bader S., Latif F., Chen J.-Y., Duh F.-M., Wei M.-H.,
 RA Albanesi J.P., Lee C.-C., Lerman M.I., Minna J.D.;
 RT "Human semaphorins A(V) and IV reside in the 3p21.3 small cell lung
 RT cancer deletion region and demonstrate distinct expression patterns."
 RL Proc. Natl. Acad. Sci. U.S.A. 93:4120-4125(1996).
 CC -!- FUNCTION: MAY PLAY A ROLE IN CELL MOTILITY AND CELL ADHESION.
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -!- TISSUE SPECIFICITY: EXPRESSED ABUNDANTLY BUT DIFFERENTIALLY IN A
 CC VARIETY OF NEURAL AND NONNEURAL TISSUES. THERE IS HIGH EXPRESSION

CC IN MAMMARY GLAND, KIDNEY, FETAL BRAIN, AND LUNG AND LOWER
 CC EXPRESSION IN HEART AND LIVER.
 CC -!- DEVELOPMENTAL STAGE: DETECTED AS EARLY AS EMBRYONIC DAY 10.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.

CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.

CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).

CC EMBL; U33920; AAC50568.1; -;
 CC EMBL; U32876; AAB18276.1; -;
 CC EMBL; AC000063; AAB46344.1; -;
 CC EMBL; U32171; AAB06011.1; -;
 CC EMBL; U32172; AAB06012.1; -;
 CC Genew; HGNC:10728; SEMA3F.

CC MIM; 601124; -;
 CC GO; GO:0005615; C:extracellular space; TAS.

CC InterPro; IPR007110; IG-like.
 CC InterPro; IPR003599; IG.

CC InterPro; IPR003006; IG_MHC.
 CC InterPro; IPR003659; plexin-like.

CC InterPro; IPR001627; Sema.
 CC Pfam; PF00047; ig; 1.

CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00409; IG; 1.

CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.

CC PROSITE; PS50835; IG-LIKE; 1.
 CC Signal; Immunoglobulin domain; Multigene family; Glycoprotein;

KW Polymorphism.
 KW SIGNAL 1 18 POTENTIAL.

FT CHAIN 19 785 SEMAPHORIN 3F.
 FT DOMAIN 272 569 SEMA.

FT DOMAIN 605 690 IG-LIKE C2-TYPE.
 FT DOMAIN 758 779 ARG/LYS-RICH (BASIC).

FT DISULFID 678 746 BY SIMILARITY.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 474 474 A -> G (in dbSNP:1046955).

FT VARIANT 503 503 L -> M.
 FT CONFLICT 153 183 /FTID=VAR_008855.

FT CONFLICT 270 270 MISSING (IN REF. 2).
 FT CONFLICT 473 473 MISSING (IN REF. 2).

FT CONFLICT 473 473 A -> S (IN REF. 2).
 SQ SEQUENCE 785 AA; 88381 MW; FE3FC796EBC1608E CRC64;

Alignment Scores:
 Pred. No.: 8,41e-05 Length: 785

Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0

DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM3F_HUMAN (1-785)

QY 1778 TGTTCCTCGCCGAGACCTTACTGTCTGGATGCT 1816

DB 558 CysCysLeuAlaArgAspProTrpCysAlaTrpAspGly 570

RESULT 19

ID SM3F_MOUSE

AC O88632; O88633; STANDARD; PRT; 785 AA.

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Semaphorin 3F precursor (Semaphorin IV) (Sema IV).
 GN SEMA3F.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=99090943; PubMed=9875739;

RA Eckhardt F., Meyerhans A.;
 RT "Cloning and expression pattern of a murine semaphorin homologous to
 H-sema IV.";

RL NeuroReport 9:3975-3979(1998).
 CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).

CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;

CC Name=B;
 CC IsoId=O88632-1; Sequence=Displayed;

CC Name=A;
 CC IsoId=O88632-2; Sequence=VSP_006042;

CC -!- TISSUE SPECIFICITY: EXPRESSED UBQUITOUSLY IN ADULTHOOD. DURING
 CC EMBRYOGENESIS, EXPRESSED IN SUBREGIONS OF THE CENTRAL NERVOUS
 CC SYSTEM AND VARIOUS OTHER TISSUES LIKE SKIN, KIDNEY, LUNG AND
 CC INTESTINE.

CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.

CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).

CC EMBL; AF080090; AAC28108.1; -;
 CC EMBL; AF080091; AAC28109.1; -;

CC MGD; MGI:1096347; Sema3f.
 CC InterPro; IPR007110; IG-like.

CC InterPro; IPR003598; IG_C2.
 CC InterPro; IPR003006; IG_MHC.

CC InterPro; IPR003659; plexin-like.
 CC InterPro; IPR001627; Sema.

CC Pfam; PF00047; ig; 1.
 CC Pfam; PF01403; Sema; 1.

CC SMART; SM00408; IGC2; 1.
 CC SMART; SM00423; PSI; 1.

CC SMART; SM00630; Sema; 1.
 CC PROSITE; PS50835; IG-LIKE; 1.

CC Signal; Immunoglobulin domain; Multigene family; Glycoprotein;
 KW Alternative splicing.

KW SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 785 SEMAPHORIN 3F.

FT DOMAIN 272 569 SEMA.
 FT DOMAIN 605 695 IG-LIKE C2-TYPE.

FT DOMAIN 758 779 ARG/LYS-RICH (BASIC).
 FT DOMAIN 715 718 POLY-PRO.

FT DISULFID 678 746 BY SIMILARITY.
 FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 153 183 Missing (in isoform A).

FT /FTID=VSP_006042
 SQ SEQUENCE 785 AA; 88493 MW; 89DBA7A08D0A272E CRC64;

Alignment Scores:
 Pred. No.: 8,41e-05 Length: 785

Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0

DB: 1 Gaps: 0

```

US-09-774-490-1 (1-2709) x SM3F_MOUSE (1-785)
QY 1778 TGTTCCTCGCCGAGACCTTACTGTGCTGGATGGT 1816
    |||||||
Db 558 CysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 570

RESULT 20
SM4C_MOUSE
ID SM4C_MOUSE STANDARD; PRT; 834 AA.
AC Q64151;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 4C precursor (Semaphorin 1) (Sema 1) (Semaphorin C-like 1)
DE (M-Sema F).
GN SEMA4C OR SEMA1 OR SEMACL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Neonatal brain;
RX MEDLINE=95385809; PubMed=7656991;
RA Inagaki S., Furuyama T., Iwahashi Y.;
RT "Identification of a member of mouse semaphorin family.";
RL FEBS Lett. 370:269-272(1995).
RN [2]
RP INTERACTION WITH GIPC.
RX MEDLINE=99253973; PubMed=10318831;
RA Wang L.-H., Kalb R.G., Strittmatter S.M.;
RT "A PDZ protein regulates the distribution of the transmembrane
semaphorin, M-SemF.";
RL J. Biol. Chem. 274:14137-14146(1999).
CC -!- SUBUNIT: INTERACTS WITH GIPC PDZ DOMAIN.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: EXPRESSED WIDELY IN THE NERVOUS TISSUES DURING
DEVELOPMENT. PREDOMINANTLY IN THE DEVELOPING BRAIN AND SPINAL
CORD.
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; S79463; AAB35184.1; -
CC PIR; S66498; S66498.
CC MGD; MGI:109252; Sema4c.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003599; Ig.
CC InterPro; IPR003006; Ig MHC.
CC InterPro; IPR003659; Plexin-like.
CC InterPro; IPR002165; Plexin_repeat.
CC InterPro; IPR001627; Sema.
CC Pfam; PF01437; PSI; 1.
CC SMART; SM00409; Sema; 1.
CC SMART; SM00423; PSI; 1.
CC SMART; SM00630; Sema; 1.
CC PROSITE; PS50835; IG LIKE; 1.
KW Signal; Transmembrane; Immunoglobulin domain; Multigene family;
KW Neurogenesis; Developmental protein; Glycoprotein.
FT SIGNAL 1 20
FT CHAIN 21 834 SEMAPHORIN 4C.
FT DOMAIN 21 664 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 665 685 POTENTIAL.

FT DOMAIN 686 834 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 53 481 SEMA.
FT DOMAIN 499 552 PSI.
FT DOMAIN 557 645 IG-LIKE C2-TYPE.
FT DOMAIN 678 687 POLY-LEU.
FT DOMAIN 709 774 PRO-RICH.
FT DOMAIN 831 834 BINDS TO GIPC.
FT DISULFID 578 628 BY SIMILARITY.
FT CARBOHYD 106 106 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 121 121 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 310 310 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 419 419 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 522 522 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 565 565 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 834 AA; 92556 MW; 6868BB5BF571482D CRC64;

Alignment Scores:
Pred. No.: 0.137 Length: 834
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM4C_MOUSE (1-834)
QY 1043 AAATGGACACATTCCTCAAAGCTCGTCTG 1072
    |||||||
Db 274 LysTrpThrPheLeuLysAlaArgLeu 283

RESULT 21
SM4D_CHICK
ID SM4D_CHICK STANDARD; PRT; 295 AA.
AC Q90655;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 4D (Collapsin-4) (COLL-4) (Fragment).
GN SEMA4D OR COLL4.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95329269; PubMed=7605628;
RA Luo Y., Shepherd I., Li J., Renzi M.J., Chang S., Raper J.A.;
RT "A family of molecules related to collapsin in the embryonic chick
nervous system.";
RL Neuron 14:1131-1140(1995).
CC -!- FUNCTION: INDUCES THE COLLAPSE AND PARALYSIS OF NEURONAL GROWTH
CONES. COULD POTENTIALLY ACT AS REPULSIVE CUES TOWARD SPECIFIC
NEURONAL POPULATIONS.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; U28242; AAA86898.1; -
CC InterPro; IPR002165; Plexin_repeat.
CC InterPro; IPR001627; Sema.
CC Pfam; PF01437; PSI; 1.
CC Pfam; PF01403; Sema; 1.
CC SMART; SM00630; Sema; 1.
KW Multigene family; Neurogenesis; Developmental protein; Glycoprotein.

```

FT NON_TER 1 1 SEMA.
 FT DOMAIN <1 255
 FT PSI 275
 FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT NON_TER 295 295
 SQ SEQUENCE 295 AA; 33524 MW; 298AA49414C352FOA CRC64;

Alignment Scores:
 Pred. No.: 1.76 Length: 295
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.01% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM4D_CHICK (1-295)

QY 1784 CTCGCCGAGACCTTACTGTCTTGG 1810
 Db 287 LeuAlaArgAspProtyrCysalaTrp 295

RESULT 22
 C1S_HUMAN STANDARD; PRT; 688 AA.
 AC P09871; Q9UCU7; Q9UCU8; Q9UCU9; Q9UCV0; Q9UCV1; Q9UCV2; Q9UCV3;
 AC Q9UCV4; Q9UCV5; Q9UM14;
 DT 01-WAR-1989 (Rel. 10, Created)
 DT 01-WAR-1989 (Rel. 10, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Complement C1s component precursor (EC 3.4.21.42) (C1 esterase).
 GN C1s.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RX SEQUENCE FROM N.A.
 RX MEDLINE=89017187; PubMed=2459702;
 RA Kusumoto H., Hirose S., Salter J.-P., Hagen F.S., Kurachi K.;
 RT "Human genes for complement components C1r and C1s in a close
 tail-to-tail arrangement.";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:7307-7311(1988).
 RN [1]
 RN SEQUENCE FROM N.A.
 RN MEDLINE=88082788; PubMed=3500856;
 RX McKinnon C.M., Carter P.E., Smyth S.J., Dunbar B., Fothergill J.E.;
 RT "Molecular cloning of cDNA for human complement component C1s. The
 complete amino acid sequence.";
 RL Eur. J. Biochem. 169:547-553(1987).
 RN [3]
 RN SEQUENCE FROM N.A.
 RN MEDLINE=8613522; PubMed=2831944;
 RX Tosi M., Duponchel C., Meo T., Julier C.;
 RT "Complete cDNA sequence of human complement C1s and close physical
 linkage of the homologous genes C1s and C1r.";
 RL Biochemistry 26:8516-8524(1987).
 RN [4]
 RN SEQUENCE OF 1-329 FROM N.A.
 RC TISSUE=Peripheral blood leukocytes;
 RX MEDLINE=99008558; PubMed=9794427;
 RA Endo Y., Takahashi M., Nakao M., Saiga H., Sekine H., Matsushita M.,
 RA Nonaka M., Fujita T.;
 RT "Two lineages of mannose-binding lectin-associated serine protease
 (MASP) in vertebrates.";
 RL J. Immunol. 161:4924-4930(1998).
 RN [5]
 RN SEQUENCE OF 291-688 FROM N.A.
 RX MEDLINE=90040704; PubMed=2553984;
 RA Tosi M., Duponchel C., Meo T., Couture-Tosi E.;
 RT "Complement genes C1r and C1s feature an intronless serine protease
 domain closely related to haptoglobin.";

RL J. Mol. Biol. 208:709-714(1989).
 RN [6]
 RN SEQUENCE OF 16-61; 168-219; 287-334 AND 384-445.
 RP MEDLINE=86164350; PubMed=3007145;
 RX Spycher S.E., Nick H., Rickli E.E.;
 RA "Human complement component C1s. Partial sequence determination of
 the heavy chain and identification of the peptide bond cleaved during
 activation.";
 RT Eur. J. Biochem. 156:49-57(1986).
 RL [7]
 RN SEQUENCE OF 438-500; 503-534; 542-601; 617-623 AND 626-656.
 RP MEDLINE=84104122; PubMed=6362661;
 RX Carter P.E., Dunbar B., Fothergill J.E.;
 RA "The serine proteinase chain of human complement component C1s.
 Cyanogen bromide cleavage and N-terminal sequences of the
 fragments.";
 RT Biochem. J. 215:565-571(1983).
 RL [8]
 RN PARTIAL SEQUENCE.
 RC TISSUE=Plasma;
 RX MEDLINE=91308095; PubMed=1854725;
 RA Illy C., Thielens N.M., Gagnon J., Arlaud G.J.;
 RT "Effect of lactoperoxidase-catalyzed iodination on the Ca(2+)-
 dependent interactions of human C1s. Location of the iodination
 sites.";
 RT Biochemistry 30:7135-7141(1991).
 RL [9]
 RN DISULFIDE BONDS.
 RP MEDLINE=91175725; PubMed=2007122;
 RX Hess D., Schaller J., Rickli E.E.;
 RA "Identification of the disulfide bonds of human complement C1s.";
 RT Biochemistry 30:2827-2833(1991).
 RL [10]
 RN PARTIAL SEQUENCE, AND 3D-STRUCTURE MODELING OF CATALYTIC DOMAIN.
 RP MEDLINE=95298736; PubMed=779774;
 RX Rossi V., Garbriaud C., Lacroix M., Ulrich J., Fontecilla-Camps J.C.,
 RA Gagnon J., Arlaud G.J.;
 RT "Structure of the catalytic region of human complement protease C1s:
 study by chemical cross-linking and three-dimensional homology
 modeling.";
 RL Biochemistry 34:7311-7321(1995).
 RN [11]
 RN DISEASE.
 RX MEDLINE=21286517; PubMed=11390518;
 RA Dragon-Durey M.-A., Quartier P., Fremaux-Bacchi V., Blouin J.,
 RA de Barace C., Prieur A.-M., Weiss L., Fridman W.-H.;
 RT "Molecular basis of a selective C1s deficiency associated with early
 onset multiple autoimmune diseases.";
 RL J. Immunol. 166:7612-7616(2001).
 CC -1- FUNCTION: C1S B CHAIN IS A SERINE PROTEASE THAT COMBINES WITH C1Q
 AND C1S TO FORM C1, THE FIRST COMPONENT OF THE CLASSICAL PATHWAY
 OF THE COMPLEMENT SYSTEM. C1R ACTIVATES C1S SO THAT IT CAN, IN
 TURN, ACTIVATE C2 AND C4.
 CC -1- CATALYTIC ACTIVITY: Cleaves component C4 to C4a and C4b (Arg-|-Ala
 bond), and component C2 to C2a and C2b (Lys(or Arg)-|-Lys bond).
 CC -1- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R
 AND S IN THE MOLAR RATION OF 1:2:2. ACTIVATED C1S IS AN DISULFIDE-
 LINKED HETERODIMER OF AN HEAVY CHAIN AND A LIGHT CHAIN.
 CC -1- DISEASE: Defects in C1s are the cause of selective C1s deficiency
 [MIM:120580]; that is associated with early onset multiple
 autoimmune diseases.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -1- SIMILARITY: Contains 2 Sushi (SCR) domains.
 CC -1- SIMILARITY: Contains 2 CUB domains.
 CC -1- SIMILARITY: Contains 1 EGF-like domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 or send an email to license@isb-sib.ch.

CC EMBL; X06596; CAA29817.1; -
 DR EMBL; J04080; AAA51852.1; -
 DR EMBL; M18767; AAA51853.1; -
 DR EMBL; AB009076; BAA86864.1; -
 DR PDB; 1ELV; 14-MAR-01.
 DR MEROPS; S01.193; -
 DR SWISS-2DPAGE; P09871; HUMAN.
 DR Genew; HGNC:1247; CIS.
 DR MIM; 120580; -
 DR GO; GO:0003816; F:complement component C1s activity; TAS.
 DR InterPro; IPR000152; Asx_hydroxyl.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000859; CUB domain.
 DR InterPro; IPR001881; EGF Ca.
 DR InterPro; IPR006209; EGF like.
 DR InterPro; IPR001254; Ser protease Try.
 DR InterPro; IPR000436; Sushi_SCR_CCF.
 DR Pfam; PF00431; CUB; 2.
 DR Pfam; PF00008; EGF; 1.
 DR Pfam; PF00084; sushi; 2.
 DR PRINTS; PR00089; trypsin; 1.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PROSITE; PS00010; ASX_HYDROXYL; 1.
 DR PROSITE; PS01180; CUB; 2.
 DR PROSITE; PS00022; EGF_1; FALSE NEG.
 DR PROSITE; PS01186; EGF_2; FALSE NEG.
 DR PROSITE; PS01187; EGF CA; 1.
 DR PROSITE; PS02040; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; FALSE NEG.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Complement pathway; Plasma; Glycoprotein; Serine protease; Hydrolase;
 KW Hydroxylation; Sushi; Repeat; Signal; EGF-like domain;
 KW Calcium-binding; Polymorphism; 3D-structure.
 FT SIGNAL 1 15
 FT CHAIN 16 437
 FT CHAIN 438 688
 FT DOMAIN 16 130
 FT DOMAIN 131 172
 FT DOMAIN 175 290
 FT DOMAIN 293 355
 FT DOMAIN 358 422
 FT DOMAIN 438 688
 FT ACT SITE 475 475
 FT ACT SITE 529 529
 FT ACT SITE 632 632
 FT MOD_RES 149 149
 FT CARBOHYD 174 174
 FT CARBOHYD 406 406
 FT DISULFID 65 83
 FT DISULFID 135 147
 FT DISULFID 143 156
 FT DISULFID 158 171
 FT DISULFID 175 202
 FT DISULFID 234 251
 FT DISULFID 294 341
 FT DISULFID 321 354
 FT DISULFID 359 403
 FT DISULFID 386 421
 FT DISULFID 425 549
 FT DISULFID 595 618
 FT DISULFID 628 659
 FT VARIANT 383 383
 FT CONFLICT 294 294
 FT CONFLICT 513 513
 FT CONFLICT 573 573
 FT CONFLICT 645 646
 SQ SEQUENCE 688 AA; 76684 MW; 85522647A4C47205 CRC64;
 Alignment Scores:
 Pred. No.: 1.64 Length: 688
 Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.03% Indels: 0
 DB: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x CIS_HUMAN (1-688)
 QY 151 TTGTGCGCCAGAGAGTTCACAAAT 125
 DB 82 LeuCysGlyGlnArgSerSerAsnAsn 90
 RESULT 23
 SM4B MOUSE
 ID SM4B MOUSE STANDARD; PRT; 782 AA.
 AC Q62179;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 4B (Semaphorin C) (Sema C) (Fragment).
 GN SEMA4B OR SEMAC OR SEMC.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NMRI; TISSUE=Brain;
 RX MEDLINE=95267431; PubMed=7748561;
 RA Puschel A.W., Adams R.H., Betz H.;
 RT "Murine semaphorin D/collapsin is a member of a diverse gene family
 and creates domains inhibitory for axonal extension.";
 RL Neuron 14:941-948 (1995).
 RP INTERACTION WITH GIPC.
 EX MEDLINE=95253973; PubMed=10318831;
 RA Wang L.-H., Kalb R.G., Strittmatter S.M.;
 RT "A PDZ protein regulates the distribution of the transmembrane
 semaphorin, M-SemF.";
 RL J. Biol. Chem. 274:14137-14146 (1999).
 CC -!- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
 SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS.
 CC -!- SUBUNIT: INTERACTS WITH GIPC PDZ DOMAIN.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE EMBRYO. LOW
 LEVELS FOUND BETWEEN DAYS 10-12. EXPRESSION PEAKS ON DAY 13 WITH
 MODERATE LEVELS FROM THEN UNTIL BIRTH.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
 or send an email to license@isb-sib.ch).

 EMBL; X85992; CAA59984.1; -
 PIR; I48746; I48746.
 MGD; MGI:107559; Sema4b.
 InterPro; IPR003659; Plexin-like.
 InterPro; IPR002165; Plexin_repeat.
 InterPro; IPR001627; Sema.
 Pfam; PF01437; PSI; 1.
 Pfam; PF01403; Sema; 1.
 SMART; SM00423; PSI; 1.
 SMART; SM00630; Sema; 1.
 Transmembrane; Immunoglobulin domain; Multigene family; Neurogenesis;
 Developmental protein; Glycoprotein.
 NON_TER 1 1
 DOMAIN <1 662 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 663 683 POTENTIAL.
 FT

FT DOMAIN 684 782 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 16 453 SEMA.
 FT DOMAIN 471 525 PSI.
 FT DOMAIN 548 608 IG-LIKE C2-TYPE.
 FT DOMAIN 703 726 PRO-RICH.
 FT DISULFID 555 601 BY SIMILARITY.
 FT CARBOHYD 12 12 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 15 15 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 42 42 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 471 471 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 526 526 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 574 574 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 639 639 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 782 AA; 36823 MW; 627A81FC8F8F7AC8 CRC64;

Alignment Scores:
 Pred. No.: 1.62 Length: 782
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.01% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM4B_MOUSE (1-782)

QY 1784 CTCGCCGAGACCTACTGTCTTGG 1810
 DB 483 LeuAlaArgAspProTyrcysAlaTrp 491

RESULT 24
 SM4B HUMAN
 ID SM4B HUMAN STANDARD; PRT: 832 AA.
 AC Q9NPR2; Q8WVP9; Q96FK5; Q9C0B8; Q9H691; Q9NPW8; Q9NPNO;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 4B precursor.
 GN SEMA4B OR KIA1745.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=21082932; PubMed=11214970;
 RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIX.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 7:347-355 (2000).
 RN [2]
 RP SEQUENCE OF 160-832 FROM N.A., AND VARIANT ALA-792.
 RA Carim L., Estivill X., Escarceller M., Sumoy L.;
 RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 358-832 FROM N.A.
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
 RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 303-832 FROM N.A., AND VARIANT ALA-792.
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Whiting J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -1- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
 CC SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS (BY
 CC SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -1- SIMILARITY: Contains 1 Sema domain.
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -1- CAUTION: Ref.3 sequence differs from that shown in position 709
 CC onward due to a frameshift.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB051532; BAB21836.1; -;
 DR EMBL; AL390080; CAB98204.1; -;
 DR EMBL; AL390081; CAB98205.1; -;
 DR EMBL; AL390082; CAB98206.1; -;
 DR EMBL; AK026133; BAB15372.1; ALT_FRAME.
 DR EMBL; BC010701; AAH10701.1; ALT_INIT.
 DR EMBL; BC017658; AAH17658.1; -;
 DR EMBL; HGNC:10730; SEMA4B.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR002165; Plexin_repeat.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01437; PSI; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00630; Sema; 1.
 DR SMART; SM00630; Sema; 1.
 DR Transmembrane; Immunoglobulin domain; Multigene family; Neurogenesis;
 DR Developmental protein; Glycoprotein; Signal; Polymorphism.
 FT SIGNAL 1 38
 FT CHAIN 39 832 SEMAPHORIN 4B.
 FT DOMAIN 39 712 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 713 733 POTENTIAL.
 FT DOMAIN 734 832 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 65 502 SEMA.
 FT DOMAIN 520 574 PSI.
 FT DOMAIN 599 658 IG-LIKE C2-TYPE.
 FT DOMAIN 753 776 PRO-RICH.
 FT DISULFID 606 651 BY SIMILARITY.
 FT CARBOHYD 64 64 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 91 91 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 405 405 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 520 520 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 625 625 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT VARIANT 792 792 S -> A.
 FT /FTID=VAR_010758.
 FT CONFLICT 160 160 N -> P (IN REF. 2).
 FT CONFLICT 565 565 E -> K (IN REF. 4; AAH10701).
 SQ SEQUENCE 832 AA; 92192 MW; 29D56C1DDDD5E1C6B CRC64;

Alignment Scores:

FT	TRANSMEM	734	754	POTENTIAL.
FT	DOMAIN	755	861	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	50	482	SEMA.
FT	DOMAIN	502	551	PSI.
FT	DOMAIN	555	636	IG-LIKE C2-TYPE.
FT	DISULFID	576	624	BY SIMILARITY.
FT	CARBOHYD	49	49	N-LINKED (GLCNAC. .)
FT	CARBOHYD	77	77	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	139	139	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	191	191	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	379	379	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	419	419	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	613	613	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	632	632	N-LINKED (GLCNAC. .) (POTENTIAL).
SQ	SEQUENCE	861 AA;	95714 MW; 533CD6D271A6D79B CRC64;	

Alignment Scores:

Pred. No.: 1.6 Length: 861

Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.01% Indels: 0

DB: I Gaps: 0

US-09-774-490-1 (1-2709) x SM4D_MOUSE (1-861)

Oy 914 GRAGATGCAAACTATCTTTTCTTC 940

Dd 227 GLuAspAspLysValTyrPhePhePhe 235

|||||

RESULT 26
SM4D_HUMAN
ID_SM4D_HUMAN STANDARD; PRT; 862 AA.
AC Q22854;
DT 30-MAY-2000 (Rel. 39, Created)
DD 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Semaphorin 4D precursor (Leukocyte activation antigen CD100) (BB18)
DE (A8) (GR3).
GN SEMA4D OR CD100.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=T-cell;
RA MEDLINE=97030273; PubMed=8876214;
RX Hall K.T., Bounsell L., Schultze J.L., Boussiotis V.A., Dorfman D.M.,
RA Cardoso A.A., Bensussan A., Nadler L.M., Freeman G.J.;
RT "Human CD100, a novel leukocyte semaphorin that promotes B-cell
RT aggregation and differentiation.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:11780-11785(1996).
CC -1- FUNCTION: MAY PLAY A FUNCTIONAL ROLE IN THE IMMUNE SYSTEM, AS WELL
CC AS IN THE NERVOUS SYSTEM. INDUCES B CELLS TO AGGREGATE AND
CC IMPROVES THEIR VIABILITY IN VITRO.
CC -1- SUBUNIT: Homodimer.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN SKELETAL MUSCLE,
CC PERIPHERAL BLOOD LYMPHOCYTES, SPLEEN, AND THYMUS AND ALSO
CC EXPRESSED AT LOWER LEVELS IN TESTES, BRAIN, KIDNEY, SMALL
CC INTESTINE, PROSTATE, HEART, PLACENTA, LUNG, AND PANCREAS BUT NOT
CC IN COLON OR LIVER.
CC -1- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1- SIMILARITY: Contains 1 sema domain.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD100 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cdi00.htm"

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U60800; AAC50810.1; -;
 DR Genew; HGNC:10732; SEMA4D.
 CC -----
 DR MIM; 601866; -;
 DR GO; GO:0006916; P:anti-apoptosis; TAS.
 DR GO; GO:0007155; P:cell adhesion; TAS.
 DR GO; GO:0006955; P:immune response; TAS.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR003006; IG_MHC.
 DR InterPro; IPR003659; plexin-like.
 DR InterPro; IPR002165; plexin_repeat.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF00047; IG; 1.
 DR Pfam; PF01437; PSI; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 KW Signal; Transmembrane; Immunoglobulin domain; Multigene family;
 KW Neurogenesis; Developmental protein; Glycoprotein.
 FT SIGNAL 1 21
 FT CHAIN 22 862
 FT DOMAIN 22 734
 FT TRANSMEM 735 755
 FT DOMAIN 756 862
 FT DOMAIN 50 482
 FT DOMAIN 502 551
 FT DOMAIN 554 636
 FT DISULFID 576 624
 FT CARBOHYD 49 49
 FT CARBOHYD 77 77
 FT CARBOHYD 139 139
 FT CARBOHYD 191 191
 FT CARBOHYD 329 329
 FT CARBOHYD 379 379
 FT CARBOHYD 419 419
 FT CARBOHYD 613 613
 FT CARBOHYD 632 632
 SQ SEQUENCE 862 AA; 96149 MW; 7B18EFA98789371 CRC64;
 Alignment Scores:
 Pred. No.: 1.6 Length: 862
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.01% Indels: 0
 DB: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x SM4D_HUMAN (1-862)
 QY 1055 TTCTCAAGCTCGTCTGATTGTCA 1081
 DB 274 PheLeuLySAlaArgLeuIleCysSer 282
 RESULT 27
 IF38_HUMAN
 ID IF38_HUMAN STANDARD; PRT; 913 AA.
 AC Q99613; Q00215;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Eukaryotic translation initiation factor 3 subunit 8 (eIF3 p110)
 DE (eIF3c).
 GN EIF38.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;

RN SEQUENCE FROM N.A.
 RP MEDLINE=97150873; PubMed=8995409;
 RX Asano K., Kinzy T.G., Merrick W.C., Hershey J.W.B.;
 RA "Conservation and diversity of eukaryotic translation initiation
 RT factor eIF3.";
 RL J. Biol. Chem. 272:1101-1109(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99425270; PubMed=10493829;
 RA Loftus B.J., Kim U.-J., Shedd V.P., Kalush F., Brandon R.,
 RA Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,
 RA Deslattes Mays A., Cao Y., Xu R.X., Kang H.-L., Mitchell S.,
 RA Eichler E.E., Harris P.C., Venter J.C., Adams M.D.;
 RT "Genome duplications and other features in 12 Mb of DNA sequence from
 RL human chromosome 16p and 16q.";
 RL Genomics 60:295-308(1999).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
 RA Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Obayashi M.,
 RA Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isegai T., Sugano S.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Berge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,
 RA Faney J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RL human and mouse cDNA sequences.";
 CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: BINDS TO THE 40S RIBOSOME AND PROMOTES THE BINDING OF
 CC METHIONYL-TRNAI AND MRNA.
 CC -1- SUBUNIT: EIF-3 IS COMPOSED OF AT LEAST 12 DIFFERENT SUBUNITS.
 CC -1- DOMAIN: CONTAINS 1 PCI DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE EIF388 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; U46025; AAD03462.1; -;
 DR EMBL; AC002544; AAC27426.1; -;
 DR EMBL; U91326; AAC27674.1; -;
 DR EMBL; AK000739; BAA91352.1; -;
 DR EMBL; BC001571; AAH01571.1; -;
 DR Genew; HGNC:3279; EIF38.
 DR GK; Q99613; -;
 DR MIM; 603916; -;
 DR GO; GO:0005852; C:eukaryotic translation initiation factor 3 . . . ; TAS.
 DR GO; GO:0003743; F:translation initiation factor activity; TAS.
 DR GO; GO:0006446; P:regulation of translational initiation; TAS.

DR InterPro; IPR000717; PCI.
DR Pfam; PF03399; PCI; 1.
DR SMART; SM00088; PINT; 1.
KW Initiation factor; Protein biosynthesis.
FT DOMAIN 164 189 ASP/GLU-RICH (ACIDIC).
FT DOMAIN 243 246 POLY-GLU.
FT DOMAIN 291 294 POLY-GLU.
FT CONFLICT 313 314 EK -> VR (IN REF. 2).
SQ SEQUENCE 913 AA; 105343 MW; CB5029F4EB51C1AA CRC64;

Alignment Scores:
Pred. No.: 1.6 Length: 913
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x IF38_HUMAN (1-913)

QY 2093 CAAGCCCTCTGCTACGTAGTCTACAA 2119
|||||
Db 648 GlnGlyLeuLeuArgSerLeuGln 656

RESULT 28

SLA1_YEAST STANDARD; PRT; 1244 AA.
AC P32790;

DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Cytoskeleton assembly control protein SLA1.
GN SLA1 OR YBL007C OR YBL0321.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;

RN [1]
RP SEQUENCE FROM N.A.

RC STRAIN=DDY 228;
RX MEDLINE=93328765; PubMed=8335689;
RA Holtzman D.A., Yang S., Drubin D.G.;
RT "Synthetic-lethal interactions identify two novel genes, SLA1 and
SLA2, that control membrane cytoskeleton assembly in Saccharomycetes
cerevisiae.";
RL J. Cell Biol. 122:635-644(1993).
RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=S288C;
RX MEDLINE=93070613; PubMed=1441753;
RA Delaveau T., Jacq C., Perea J.;
RT "Sequence of a 12.7 kb segment of yeast chromosome II identifies a
PDR-like gene and several new open reading frames.";
RL Yeast 8:761-768(1992).

CC -!- FUNCTION: ESSENTIAL FOR THE PROPER FORMATION OF THE CORTICAL

CC ACTIN CYTOSKELETON. INVOLVED IN CONTROLLING THE SIZE OF CORTICAL
CC PATCHES PERHAPS BY REGULATING THE NUCLEATION OF FILAMENTS AT THE
CC CORTEX.

CC -!- SIMILARITY: SOME. TO SEA URCHIN BINDIN.

CC -!- SIMILARITY: Contains 3 SH3 domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; 222810; CAA80463.1; -.

CC EMBL; 235768; CAA84826.1; -.

CC EMBL; 547695; AAB23985.1; -.

CC PIR; S25327; S25327.

DR HSSP; Q03526; LAWJ.
DR SGD; S0000103; SLA1.
DR GO; GO:0005857; C:actin cortical patch (sensu Saccharomycetes); IDA.
DR GO; GO:0007015; P:actin filament organization; IMP.
DR GO; GO:0007047; P:cell wall organization and biogenesis; IMP.
DR GO; GO:0006897; P:endocytosis; IMP.
DR GO; GO:0007121; P:polar budding; IMP.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00018; SH3; 3.
DR Pfam; PF03983; SHD1; 1.
DR PRINTS; PR00452; SH3DOMAIN.
DR ProDom; PD000066; SH3; 3.
DR SMART; SM00326; SH3; 3.
DR PROSITE; PS00002; SH3; 3.
KW Cytoskeleton; Actin-binding; SH3 domain; Repeat.
FT DOMAIN 8 69
FT DOMAIN 70 132
FT DOMAIN 353 415
FT DOMAIN 868 1205

16 X 7 AA APPROXIMATE REPEATS OF
T-G-G-A-M-N-P.

FT REPEAT 868 874
FT REPEAT 877 883
FT REPEAT 887 893
FT REPEAT 923 929
FT REPEAT 945 951
FT REPEAT 1003 1009
FT REPEAT 1020 1026
FT REPEAT 1031 1037
FT REPEAT 1048 1054
FT REPEAT 1065 1071
FT REPEAT 1084 1090
FT REPEAT 1129 1135
FT REPEAT 1155 1161
FT REPEAT 1170 1176
FT REPEAT 1185 1191
FT REPEAT 1200 1206
SQ SEQUENCE 1244 AA; 135848 MW; 7FD85AA776407624 CRC64;

Alignment Scores:

Pred. No.: 1.55 Length: 1244
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SLA1_YEAST (1-1244)

QY 380 CTTTGGATGAGAACGAGTAGGCTG 406
|||||

Db 607 LeuLeuAspGluGluArgSerArgLeu 615
|||||

RESULT 29

RYR3_HUMAN

ID RYR3_HUMAN STANDARD; PRT; 4870 AA.

AC Q15413; Q15175; Q15412;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Ryanodine receptor 3 (Brain-type ryanodine receptor) (RYR3) (RYR-3)
(Brain ryanodine receptor-calcium release channel).

GN RYR3 OR HBRR.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.; ALTERNATIVE SPLICING, AND VARIANTS.

RC TISSUE=Fetal brain;

RX MEDLINE=98175492; PubMed=9515741;

RA Leeb T., Brenig B.;

RT "cDNA cloning and sequencing of the human ryanodine receptor type 3
(RYR3) reveals a novel alternative splice site in the RYR3 gene.";

```

RL FEBS Lett. 423:367-370(1998).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND TISSUE SPECIFICITY.
RX TISSUE=Brain;
RX MEDLINE=98055474; PubMed=9395096;
RA Nakashima Y., Nishimura S., Maeda A., Barsomian E.L., Hakamata Y.,
RA Nakai J., Allen P.D., Imoto K., Kita T.;
RT "Molecular cloning and characterization of a human brain ryanodine
RT receptor.";
RL FEBS Lett. 417:157-162(1997).
RN [3]
RP SEQUENCE OF 520-660 FROM N.A., AND TISSUE SPECIFICITY.
RX TISSUE=Skeletal muscle;
RX MEDLINE=98268728; PubMed=9607712;
RA Martin C., Chapman K.E., Seckl J.R., Ashley R.H.;
RT "Partial cloning and differential expression of ryanodine
RT receptor/calcium-release channel genes in human tissues including the
RT hippocampus and cerebellum.";
RL Neuroscience 85:205-216(1998).
RN [4]
RP SEQUENCE OF 3943-4870 FROM N.A.
RX TISSUE=T-cell;
RX MEDLINE=95010709; PubMed=7523185;
RA Hakamata Y., Nishimura S., Nakai J., Nakashima Y., Kita T., Imoto K.;
RT "Involvement of the brain type of ryanodine receptor in T-cell
RT proliferation.";
RL FEBS Lett. 352:206-210(1994).
RN [5]
RP SEQUENCE OF 4644-4842 FROM N.A.
RX TISSUE=Cervical carcinoma, and Hepatoma;
RX MEDLINE=94102751; PubMed=8276408;
RA Sorrentino V., Giannini G., Malzac P., Mattei M.-G.;
RT "Localization of a novel ryanodine receptor gene (RYR3) to human
RT chromosome 15q14-q15 by in situ hybridization.";
RL Genomics 18:163-165(1993).
RN [6]
RP SEQUENCE OF 4652-4803 FROM N.A.
RX TISSUE=Myometrium;
RX MEDLINE=96032536; PubMed=7556644;
RA Lynn S., Morgan J.M., Lamb H.K., Meisner G., Gillespie J.I.;
RT "Isolation and partial cloning of ryanodine-sensitive Ca2+ release
RT channel protein isoforms from human myometrial smooth muscle.";
RL FEBS Lett. 372:6-12(1995).
CC -!- FUNCTION: Communication between transverse-tubules and
CC sarcoplasmic reticulum. Contraction of skeletal muscle is
CC triggered by release of calcium ions from SR following
CC depolarization of T-tubules (By similarity).
CC -!- SUBUNIT: Homotetramer (Potential).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1;
CC IsoId=Q15413-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q15413-2; Sequence=VSP_005954;
CC Name=3;
CC IsoId=Q15413-3; Sequence=VSP_005955, VSP_005956;
CC -!- TISSUE SPECIFICITY: Brain, skeletal muscle, placenta and possibly
CC liver and kidney. In brain, highest levels are found in the
CC cerebellum, hippocampus, caudate nucleus and amygdala, with lower
CC levels in the corpus callosum, substantia nigra and thalamus.
CC -!- MISCELLANEOUS: The calcium release channel is modulated by calcium
CC ions, magnesium ions, ATP and calmodulin.
CC -!- MISCELLANEOUS: The calcium release channel activity resides in the
CC C-terminal region while the remaining part of the protein
CC constitutes the 'foot' structure spanning the junctional gap
CC between the SR and the T-tubule. It is possible that the foot
CC structure interacts with the cytoplasmic region of the
CC dihydropyridine receptor.
CC -!- MISCELLANEOUS: Ryanodine is an alkaloid that binds to the Ca-
CC release channel in junctional SR and modulates its activity.
CC -!- SIMILARITY: BELONGS TO THE RYANODINE RECEPTOR FAMILY.
CC -!- SIMILARITY: Contains 3 SPRY domains.

```

```

CC CC -!- CAUTION: Ref.2 sequence differs from that shown due to probable
CC frameshift errors at position 742-766.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ001515; BAA04798.1; -
CC EMBL; AB001025; BAA23795.1; -
CC EMBL; AJ002512; CAA05503.1; -
CC EMBL; X74269; CAA52326.1; -
CC EMBL; X74270; CAA52327.1; -
CC F01; S66631; S66631.
CC Gene; HGNC:10485; RYR3.
CC MIM; 180903; -
CC GO; GO:0015278; F:calcium-release channel activity; TAS.
CC GO; GO:0006816; P:calcium ion transport; TAS.
CC InterPro; IPR000699; Ca-rel_channel.
CC InterPro; IPR001682; Ca/Na_pore.
CC InterPro; IPR002048; EF-hand.
CC InterPro; IPR005821; Ion_trans.
CC InterPro; IPR003608; MIR_.
CC InterPro; IPR001215; Ryanodn_receptor.
CC InterPro; IPR003032; RYR.
CC InterPro; IPR003877; SPRY_receptor.
CC Pfam; PF00036; ehand; 1.
CC Pfam; PF00520; Ion_trans; 1.
CC Pfam; PF02815; MIR_4.
CC Pfam; PF01365; RYDR_IPR; 2.
CC Pfam; PF02026; RYR; 4.
CC Pfam; PF00622; SPRY; 3.
CC PRINTS; PR00795; RYANODINER.
CC SMART; SM00472; MIR; 4.
CC SMART; SM00449; SPRY; 3.
CC Receptor; Transmembrane; Ionic channel; Calcium channel; Repeat;
CC Glycoprotein; Multigene family; Polymorphism; Alternative splicing.
CC DOMAIN 1 2984 CYTOPLASMIC (BY SIMILARITY).
CC TRANSMEM 2985 3005 M' (POTENTIAL).
CC TRANSMEM 3049 3067 M' (POTENTIAL).
CC TRANSMEM 3836 3855 M1 (POTENTIAL).
CC TRANSMEM 3874 3892 M2 (POTENTIAL).
CC TRANSMEM 4132 4153 M3 (POTENTIAL).
CC TRANSMEM 4195 4215 M4 (POTENTIAL).
CC TRANSMEM 4410 4431 M5 (POTENTIAL).
CC TRANSMEM 4482 4505 M6 (POTENTIAL).
CC TRANSMEM 4622 4642 M7 (POTENTIAL).
CC TRANSMEM 4670 4689 M8 (POTENTIAL).
CC TRANSMEM 4712 4731 M9 (POTENTIAL).
CC TRANSMEM 4756 4770 M10 (POTENTIAL).
CC DOMAIN 657 795 SPRY 1.
CC DOMAIN 1084 1207 SPRY 2.
CC DOMAIN 1325 1465 SPRY 3.
CC DOMAIN 840 2820 4 X APPROXIMATE REPEATS.
CC REPEAT 840 953 1.
CC REPEAT 954 1068 2.
CC REPEAT 2589 2707 3.
CC REPEAT 2708 2820 4.
CC CARBOHYD 3276 3276 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 3760 3760 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 3801 3801 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 4000 4000 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 4007 4007 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 4697 4697 N-LINKED (GLCNAC. .) (POTENTIAL).
CC VARSPLIC 3337 3341 Missing (in isoform 2).
CC /FTid=VSP_005954.
CC DSS -> GMW (in isoform 3).
CC /FTid=VSP_005955.
CC VARSPLIC 3857 3859 Missing (in isoform 3).
CC /FTid=VSP_005956.

```


inherited disease with a relatively early onset and a mortality rate of approximately 30% by the age of 30 years. Phenotypically, it is characterized by salvos of bidirectional and polymorphic ventricular tachycardias in response to vigorous exercise, with no structural evidence of myocardial disease.

-!- MISCELLANEOUS: The calcium release channel is modulated by calcium ions, magnesium ions, ATP and calmodulin.

-!- MISCELLANEOUS: The calcium release channel activity resides in the C-terminal region while the remaining part of the protein constitutes the 'foot' structure spanning the junctional gap between the SR and the T-tubule. It is possible that the foot structure interacts with the cytoplasmic region of the dihydropyridine receptor.

-!- MISCELLANEOUS: Ryanodine is an alkaloid that binds to the Ca-release channel in junctional SR and modulates its activity.

-!- SIMILARITY: BELONGS TO THE RYANODINE RECEPTOR FAMILY.

-!- SIMILARITY: Contains 3 SPRY domains.

This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

EMBL; X98330; CAA66975.1; -
 EMBL; Y08218; CAA69395.1; -
 EMBL; X91869; CAA62975.1; -
 EMBL; A0002511; CAA05502.1; -
 PIR; S72269; S72269.
 Genew; HGNC:10484; RYR2.
 MIM; 180902; -
 MIM; 192605; -
 MIM; 600996; -
 MIM; 604772; -
 GO; GO:0015278; P:calcium-release channel activity; TAS.
 GO; GO:0006936; P:muscle contraction; TAS.
 GO; GO:0008016; P:regulation of heart; TAS.
 InterPro; IPR000699; Ca-rel_channel.
 InterPro; IPR001882; Ca/Na_Pore.
 InterPro; IPR002048; EF-hand.
 InterPro; IPR005821; Ion_trans.
 InterPro; IPR003608; MIR.
 InterPro; IPR001215; Ryanodn_receptor.
 InterPro; IPR003032; RYR.
 InterPro; IPR003877; SPRY_receptor.
 Pfam; PF00036; ehand; 2.
 Pfam; PF00520; ion_trans; 1.
 Pfam; PF02815; MIR; 4.
 Pfam; PF01365; RYDR_ITPR; 2.
 Pfam; PF02026; RYR; 4.
 Pfam; PF00622; SPRY; 3.
 PRINTS; PR00795; RYANODINER.
 SMART; SM00054; EFh; 2.
 SMART; SM00472; MIR; 4.
 SMART; SM00449; SPRY; 3.
 Ionic channel; Receptor; Calcium channel; Calmodulin-binding; Repeat; Transmembrane; Glycoprotein; Phosphorylation; Alternative splicing; Polymorphism; Disease mutation; Cardiomyopathy; Multigene family. CYTOPLASMIC (BY SIMILARITY).

DOMAIN 1 3089
 TRANSMEM 3090 3109 M' (POTENTIAL).
 TRANSMEM 3153 3171 M'' (POTENTIAL).
 TRANSMEM 3940 3959 M1 (POTENTIAL).
 TRANSMEM 3978 3995 M2 (POTENTIAL).
 TRANSMEM 4233 4256 M3 (POTENTIAL).
 TRANSMEM 4294 4314 M4 (POTENTIAL).
 TRANSMEM 4500 4520 M5 (POTENTIAL).
 TRANSMEM 4578 4600 M6 (POTENTIAL).
 TRANSMEM 4720 4740 M7 (POTENTIAL).
 TRANSMEM 4768 4786 M8 (POTENTIAL).
 TRANSMEM 4810 4827 M9 (POTENTIAL).
 TRANSMEM 4845 4867 M10 (POTENTIAL).

FT DOMAIN 670 808
 FT DOMAIN 1098 1221
 FT DOMAIN 1423 1561
 FT DOMAIN 4414 4455
 FT REPEAT 853 2925
 FT REPEAT 967 966
 FT REPEAT 967 1080
 FT REPEAT 2692 2810
 FT REPEAT 2812 2925
 FT BINDING 2618 3015
 FT BINDING 2774 2806
 FT BINDING 2876 2897
 FT BINDING 2997 3015
 FT MOD_RES 2808 2808
 FT CARBOHYD 3427 3427
 FT CARBOHYD 3536 3536
 FT CARBOHYD 3770 3770
 FT CARBOHYD 3864 3864
 FT CARBOHYD 3865 3865
 FT CARBOHYD 3905 3905
 FT CARBOHYD 4075 4075
 FT CARBOHYD 4104 4104
 FT CARBOHYD 4111 4111
 FT CARBOHYD 4794 4794
 FT VARSPLIC 3715 3715
 FT VARIANT 433 433

Alignment Scores:
 Pred. No.: 1.38 Length: 4967
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.03% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x RYR2_HUMAN (1-4967)

Qy 67 CGTGGTGGACACCCCTGACGTGGCA 41.
 |||||
 Db 2235 ArgGlySerThrProLeuAspValAla 2243

RESULT 31
 RYR2_RABIT
 ID_RYR2_RABIT STANDARD; PRT; 4969 AA.
 AC P30957;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ryanodine receptor 2 (Cardiac muscle-type ryanodine receptor) (RYR2)
 DE (RYR-2) (Cardiac muscle ryanodine receptor-calcium release channel).
 GN RYR2.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart muscle;
 RX MEDLINE=90337947; PubMed=2380170;
 RA Otsu K., Willard H.F., Khanna V.K., Zorzato F., Green N.M.,
 MacLennan D.H.;
 RT "Molecular cloning of cDNA encoding the Ca2+ release channel
 RT (ryanodine receptor) of rabbit cardiac muscle sarcoplasmic
 RT reticulum.";
 RL J. Biol. Chem. 265:13472-13483 (1990).
 RN [2]
 RP PHOSPHORYLATION OF SER-2809.
 RX MEDLINE=91250425; PubMed=1645727;
 RA Witcher D.R., Kovacs R.J., Schulman H., Cefali D.C., Jones L.R.;
 RT "Unique phosphorylation site on the cardiac ryanodine receptor
 RT regulates calcium channel activity.";

RL J. Biol. Chem. 266:11144-11152(1991).

CC -!- FUNCTION: Communication between transverse-tubules and sarcolemmal reticulum. Contraction of cardiac muscle is triggered by release of calcium ions from SR following depolarization of T-tubules.

CC -!- SUBUNIT: Homotetramer (Potential).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).

CC -!- TISSUE SPECIFICITY: HEART AND BRAIN.

CC -!- MISCELLANEOUS: The calcium release channel is modulated by calcium ions, magnesium ions, ATP and calmodulin.

CC -!- MISCELLANEOUS: The calcium release channel activity resides in the C-terminal region while the remaining part of the protein constitutes the 'foot' structure spanning the junctional gap between the SR and the T-tubule. It is possible that the foot structure interacts with the cytoplasmic region of the dihydropyridine receptor.

CC -!- MISCELLANEOUS: Ryanodine is an alkaloid that binds to the Ca-release channel in junctional SR and modulates its activity.

CC -!- SIMILARITY: BELONGS TO THE RYANODINE RECEPTOR FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M59743; AAA31179.1; -.

DR PIR; A37113; A37113.

DR InterPro; IPR000899; Ca-rel_channel.

DR InterPro; IPR001682; Ca/Na_pore.

DR InterPro; IPR002048; EF-hand.

DR InterPro; IPR005821; Ion trans.

DR InterPro; IPR003608; MIR.

DR InterPro; IPR001215; Ryanodn_receptor.

DR InterPro; IPR003032; RYR.

DR InterPro; IPR003877; SPRY_receptor.

DR Pfam; PF00036; EFhand; 2.

DR Pfam; PF00520; ion trans; 1.

DR Pfam; PF02815; MIR; 4.

DR Pfam; PF01365; RYDR_ITPR; 2.

DR Pfam; PF02026; RYR; 4.

DR Pfam; PF00622; SPRY; 3.

DR PRINTS; PR00795; RYANODINER.

DR SMART; SM00054; Efh; 2.

DR SMART; SM00472; MIR; 4.

DR SMART; SM00449; SPRY; 3.

KW Receptor; Transmembrane; Ionic channel; Calcium channel; Repeat; Phosphorylation; Glycoprotein; Multigene family; Calmodulin-binding.

FT DOMAIN 1 3090

FT TRANSMEM 3091 3110

FT M' (POTENTIAL).

FT M1 (POTENTIAL).

FT M2 (POTENTIAL).

FT M3 (POTENTIAL).

FT M4 (POTENTIAL).

FT M5 (POTENTIAL).

FT M6 (POTENTIAL).

FT M7 (POTENTIAL).

FT M8 (POTENTIAL).

FT M9 (POTENTIAL).

FT M10 (POTENTIAL).

FT SPRY 1.

FT SPRY 2.

FT SPRY 3.

FT DOMAIN 1221

FT DOMAIN 1423

FT DOMAIN 853

FT REPEAT 853 966

FT REPEAT 967 1080

FT REPEAT 2693 2811

FT REPEAT 2813 2926

FT BINDING 2619 3016

FT CALMODULIN (POTENTIAL).

FT BINDING 2877 2898 CALMODULIN (POTENTIAL).

FT BINDING 2998 3016 CALMODULIN (POTENTIAL).

FT MOD RES 2809 2809 PHOSPHORYLATION (BY CAMK).

FT CARBOHYD 4105 4105 N-LINKED (GLCNAC...) (POTENTIAL).

FT CARBOHYD 4796 4796 N-LINKED (GLCNAC...) (POTENTIAL).

SQ SEQUENCE 4969 AA; 565060 MW; PF6E0684B974BB4D CRC64;

Alignment Scores:

Pred. No.: 1.38 Length: 4969

Score: 9.00 Matches: 9

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.03% Indels: 0

DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x RYR2_RABIT (1-4969)

QY 67 CGTGGTCGACACCCCTCGAGTGGCA 41

Db 2235 ArgGlySerThrProLeuAspValala 2243

RESULT 32

PND1_ECOLI

ID_PND1_ECOLI STANDARD; PRT; 50 AA.

AC P11902;

DT 01-OCT-1989 (Rel. 12, Created)

DT 01-OCT-1989 (Rel. 12, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE PND protein.

GN PND.

OS Escherichia coli.

OG Plasmid R16.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;

OC Enterobacteriaceae; Escherichia.

OX NCBI_TaxID:562;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89150247; PubMed=2465777;

RA Sakikawa T., Akimoto S., Ohnishi Y.;

RT "The pnd gene in E. coli plasmid R16: nucleotide sequence and gene expression leading to cell Mg2+ release and stable RNA degradation.";

RL Biochim. Biophys. Acta 1087:158-166(1989).

CC -!- FUNCTION: WHEN OVEREXPRESSED KILL THE CELLS FROM THE INSIDE BY INTERFERING WITH A VITAL FUNCTION IN THE CELL MEMBRANE.

CC -!- FUNCTION: Involved in cell Mg(2+) release and stable RNA degradation.

CC -!- INDUCTION: IN THE PRESENCE OF RIFAMPICIN AT 42 DEGREES CELSIUS.

CC -!- SIMILARITY: BELONGS TO THE HOK/GEF FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; X12833; CAA31320.1; -.

DR PIR; S02763; S02763.

DR InterPro; IPR000021; Hok/gef_toxin.

DR Pfam; PF01848; HOK_GEF; 1.

DR PRINTS; PR00281; HOKGEFTOXIC.

DR PROSITE; PS00556; HOK_GEF; 1.

KW Transmembrane; Plasmid.

FT TRANSMEM 5 25

FT SEQUENCE 50 AA; 5796 MW; F4757270A3F21792 CRC64;

Alignment Scores:

Pred. No.: 24.2 Length: 50

Score: 8.00 Matches: 8

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 0.89% Indels: 0

```

DB: 1 Gaps: 0
US-09-774-490-1 (1-2709) x PND1_ECOLI (1-50)
QY 1374 CAAGGACCTTCCTGATGATGTTA 1397
Db 3 GlnArgThrPheLeuMetMetLeu 10

RESULT 33
PND2_ECOLI STANDARD; PRT; 50 AA.
AC P16477;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE PND protein.
GN PND.
OS Escherichia coli.
OG Plasmid Inci1 R483.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88174457; PubMed=3328043;
RA Ono K., Akimoto S., Ohnishi Y.;
RT "Nucleotide sequence of the pnd gene in plasmid R483 and role of the
RT pnd gene product in plasmolysis.";
RL Microbiol. Immunol. 31:1071-1083(1987).
CC -!- FUNCTION: WHEN OVEREXPRESSED KILL THE CELLS FROM THE INSIDE BY
CC INTERFERING WITH A VITAL FUNCTION IN THE CELL MEMBRANE.
CC -!- FUNCTION: Involved in cell Mg(2+) release and stable RNA
CC degradation.
CC -!- SIMILARITY: BELONGS TO THE HOK/GEF FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D00364; BAA00270.1; -.
DR PIR; J00071; J00071.
DR InterPro; IPR000021; Hok/gef_toxin.
DR Pfam; PF01848; HOK_GEF; 1.
DR PRINTS; PR00281; HOKGEFTOXIC.
DR PROSITE; PS00556; HOK_GEF; 1.
KW Transmembrane, plasmid.
FT TRANSMEM 5 25 POTENTIAL.
SQ SEQUENCE 50 AA; 5775 MW; 2DA904DCBE81AFA9 CRC64;

Alignment Scores:
Pred. No.: 24.2 Length: 50
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PND2_ECOLI (1-50)
QY 1374 CAAGGACCTTCCTGATGATGTTA 1397
Db 3 GlnArgThrPheLeuMetMetLeu 10

RESULT 34
MTRB_BACSU STANDARD; PRT; 75 AA.
AC P19466;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)

```

```

DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Transcription attenuation protein mtrB (Tryptophan RNA-binding
DE attenuator protein) (Trp RNA-binding attenuation protein) (TRAP).
GN MTRB.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91062353; PubMed=2123343;
RA Gollnick P., Ishino S., Kuroda M.I., Henner D.J., Yanofsky C.;
RT "The mtr locus is a two-gene operon required for transcription
RT attenuation in the trp operon of Bacillus subtilis.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:8726-8730(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RC MEDLINE=98044033; PubMed=9384377;
RA Kuznetsov F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Chim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Haisch J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portelle D., Porwollik S., Prescott A.M.,
RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadale Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorek N., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tanakoshi A., Tanaka T., Terpatra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis.";
RL Nature 390:249-256(1997).
RN [3]
RP FUNCTION.
RX MEDLINE=92202128; PubMed=1551827;
RA Babinzke P., Gollnick P., Yanofsky C.;
RT "The mtrAB operon of Bacillus subtilis encodes GTP cyclohydrolase I
RT (McrA), an enzyme involved in folic acid biosynthesis, and MtrB, a
RT regulator of tryptophan biosynthesis.";
RL J. Bacteriol. 174:2059-2064(1992).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
RX MEDLINE=95231646; PubMed=7715723;
RA Antson A.A., Otridge J., Brzozowski A.M., Dodson E.J., Dodson G.G.,
RA Wilson K.S., Smith T.M., Van M., Kurecki T., Gollnick P.;
RT "The structure of trp RNA-binding attenuation protein.";
RL Nature 374:693-700(1995).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX MEDLINE=95055737; PubMed=7525975;
RA Antson A.A., Brzozowski A.M., Dodson E.J., Dauter Z., Wilson K.S.,
RA Kurecki T., Otridge J., Gollnick P.;
RT "11-fold symmetry of the trp RNA-binding attenuation protein (TRAP)
RT from Bacillus subtilis determined by X-ray analysis.";
RL J. Mol. Biol. 244:1-5(1994).
CC -!- FUNCTION: REQUIRED FOR TRANSCRIPTION ATTENUATION CONTROL IN THE

```

TRP OPERON. THIS TRANS-ACTING FACTOR SEEMS TO RECOGNIZE A 10 BASES NUCLEOTIDE SEQUENCE IN THE TRP LEADER TRANSCRIPT CAUSING TRANSCRIPTION TERMINATION. BINDS THE LEADER RNA ONLY IN PRESENCE OF L-TRYPTOPHAN.

-!- SUBUNIT: OLIGOMER OF 11 IDENTICAL SUBUNITS ARRANGED IN DOUGHNUT-LIKE STRUCTURE.

-!- SIMILARITY: WITH REGA FROM PHAGE T4.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M37320; AAA22616.1; --
 EMBL; M80245; AAA20853.1; --
 EMBL; Z99115; CAB14193.1; --
 PIR; B38256; B38256.
 PDB; 1WAP; 03-JUN-95.
 Subtilist; BG10278; mtrB.
 InterPro; IPR000824; TrpBP.
 Pfam; PF02081; TrpBP; 1.
 PRINTS; PR00687; TRPRNAP.
 ProDom; PD027518; TrpBP; 1.
 Transcription regulation; RNA-binding; 3D-structure;
 Complete proteome.

STRAND 9 14
 STRAND 19 25
 STRAND 32 38
 TURN 40 41
 STRAND 43 47
 STRAND 52 58
 STRAND 61 65
 TURN 66 67
 STRAND 68 72
 SEQUENCE 75 AA; 8328 MW; 34C173A8D15A4B31 CRC64;

Alignment Scores:
 Pred. No.: 23.3 Length: 75
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservations: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x MTRB_BACSU (1-75)

Qy 2034 GCGGAATGAGCGGAAAGAG 2057
 |||||
 Db 68 GlyGluMetLysSerGluLysLys 75

RESULT 35
 VSPC_TRIGA
 ID_VSPC_TRIGA STANDARD; PRT; 257 AA.
 AC O13062;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Venom serine proteinase 2C precursor (EC 3.4.21.-).
 GN TLG2C.
 OS Trimeresurus gramineus (Indian green tree viper) (Green habu snake).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidodermata; Squamata; Scleroglossa; Serpentes; Colubroidea;
 OC Viperidae; Crotalinae; Trimeresurus.
 OX NCBI_TaxID=8767;
 RN [1]
 RC SEQUENCE FROM N.A.
 RX TISSUE=Venom gland;
 RA MEDLINE=97096898; PubMed=8941719;
 RA Doshinaru M., Ogawa T., Nakashima K., Nobuhisa I., Chijiwa T.,
 RA Shimonigashi Y., Fukumaki Y., Niwa M., Yamashina I., Hattori S.,

Ohno M.;
 "Accelerated evolution of crotalinae snake venom gland serine proteases.";
 FEBS Lett. 397:83-88(1996).
 CC -!- FUNCTION: Thrombin-like snake venom serine protease.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Expressed by the venom gland.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. SNAKE VENOM SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; D67084; BAA19982.1; --
 HSPF; P00763; LDPO.
 MEROPS; S01.185; --
 InterPro; IPR001314; Chymotrypsin.
 Pfam; PF00089; trypsin; 1.
 PRINTS; PR00722; CHYMOTRYPSIN.
 SMART; SM00020; Tryp_Spc; 1.
 PROSITE; PS00240; TRYPSIN_DOM; 1.
 PROSITE; PS00134; TRYPSIN_HIS; 1.
 PROSITE; PS00135; TRYPSIN_SER; 1.
 Hydrolase; Serine protease; Glycoprotein; Zymogen; Signal.
 SIGNAL 1 18
 PROPEP 19 24
 CHAIN 25 257
 ACT_SITE 64 64
 ACT_SITE 109 109
 ACT_SITE 203 203
 DISULFID 31 162
 DISULFID 49 65
 DISULFID 97 255
 DISULFID 141 209
 DISULFID 173 188
 DISULFID 199 224
 CARBOHYD 116 116
 CARBOHYD 120 120
 CARBOHYD 121 121
 SEQUENCE 257 AA; 28042 MW; 2F50B6947CB33AB1 CRC64;

Alignment Scores:
 Pred. No.: 21 Length: 257
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservations: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.93% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x VSPC_TRIGA (1-257)

Qy 1998 TTCGACTGCATTCGAAAAATGTG 1975
 |||||
 Db 74 PheGlyLeuHisSerLysAsnVal 81

RESULT 36
 GRBB_BACSU
 ID_GRBB_BACSU STANDARD; PRT; 368 AA.
 AC P39570;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Spore germination protein B2.
 GN GERBB.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=94282292; PubMed=8012571;
 RA Corfe B.M., Sammons R.L., Smith D.A., Maueel C.;
 RT "The gerB region of the Bacillus subtilis 168 chromosome encodes a
 RL homologue of the gerA spore germination operon";
 RN Microbiology 140:471-478(1994).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Mosser I., Albertini A.M., Alloni G.,
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borriass R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Entian K.D., Fabbro C., Fabbro C., Ferrich S., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadate Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RL subtilis";
 Nature 390:249-256(1997).
 CC -!- FUNCTION: INVOLVED IN THE RESPONSE TO THE GERMINATIVE MIXTURE OF
 CC L-ASPARAGINE, GLUCOSE, FRUCTOSE AND POTASSIUM IONS (AGFK). COULD
 CC BE AN AMINO ACID TRANSPORTER. CANNOT STIMULATE GERMINATION IN THE
 CC ABSENCE OF ACID AND GERK GENE PRODUCTS (FRUCTOSE AND GLUCOSE
 CC RECEPTORS RESPECTIVELY).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN THE FORESPORE COMPARTMENT OF THE
 CC DEVELOPING SPORANGIUM.
 CC -!- SIMILARITY: BELONGS TO THE ARCD/CADB/GERAB/LYSI FAMILY OF
 CC PERMEASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L16960; AAA22467.1; -;
 CC PIR; 299122; CAB15598.1; -;
 CC PIR; I39856; I39856.
 CC SubtilList; BG10641; gerB.
 CC InterPro; IPR002293; AA/rel_permease1.
 CC InterPro; IPR002422; AA/rel_permease2.
 CC InterPro; IPR002091; ArAA_permease.
 CC InterPro; IPR004761; Spore_permease.
 CC Pfam; PF03845; Spore_permease; 1.
 CC PRINTS; PR00186; ARONAPRMEASE.
 CC TIGRFAMs; TIGR00912; 2A0309; 1.
 KW Germination; Transport; Amino-acid transport; Transmembrane;

KW Complete proteome. 30 POTENTIAL.
 FT TRANSMEM 10 POTENTIAL.
 FT TRANSMEM 43 POTENTIAL.
 FT TRANSMEM 82 POTENTIAL.
 FT TRANSMEM 120 POTENTIAL.
 FT TRANSMEM 145 POTENTIAL.
 FT TRANSMEM 187 POTENTIAL.
 FT TRANSMEM 217 POTENTIAL.
 FT TRANSMEM 247 POTENTIAL.
 FT TRANSMEM 282 POTENTIAL.
 FT TRANSMEM 308 POTENTIAL.
 FT TRANSMEM 338 POTENTIAL.
 SQ SEQUENCE 368 AA; 41709 MW; AB6B769512B16806 CRC64;
 Alignment Scores: 20.3 Length: 368
 Pred. No.: 8.00 Matches: 8
 Score: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x GRBB_BACSU (1-368)
 QY 2481 GTCTCTCTGTTTCTACCTTTCTTA 2458
 DB 56 ValLeuLeuPheLeuProPheLeu 63
 RESULT 37
 ID SPB8 HUMAN STANDARD; PRT; 374 AA.
 AC P50452;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytoplasmic antipeptidase 2 (CAP2) (CAP-2) (Protease inhibitor 8)
 DE (Serpin B8).
 GN SERPINB8 OR P18.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=96102039; PubMed=8530382;
 RA Sprecher C.A., Morgenstern K.A., Mathewes S., Dahlen J.R.,
 RA Schrader S.K., Foster D.C., Kisiel W.;
 RT "Molecular cloning, expression, and partial characterization of two
 RT novel members of the ovalbumin family of serine proteinase
 RT inhibitors";
 RL J. Biol. Chem. 270:29854-29861(1995).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: BELONGS TO THE SERPIN FAMILY. OV-SERPIN SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L40377; AAC41939.1; -;
 CC PIR; A59273; A59273.
 CC HSSP; P05120; 1BY7.
 CC Genew; HGNC:8952; SERPINB8.
 CC MIM; 601697; -;
 CC CO; GO:0005829; Cytosol; TAS.
 CC GO; GO:0005515; F:protein binding activity; TAS.
 CC GO; GO:0004868; F:serpin; TAS.
 CC InterPro; IPR000215; Serpin.
 CC Pfam; PF00079; serpin; 1.

```

DR SMART; SMO0093; SERPIN; 1.
DR PROSITE; PS00284; SERPIN; 1.
KW Serpin; Serine protease inhibitor.
FT ACT SITE 339 340 REACTIVE BOND (BY SIMILARITY).
SQ SEQUENCE 374 AA; 42786 MW; E855A033519AC60E CRC64;

Alignment Scores:
Pred. No.: 20.3 Length: 374
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SPB8_HUMAN (1-374)
OY 1590 TTCCTAAGCAGACTGGTATGATT 1613
Db 282 PheLeuArgA-gleuGlyMetile 289

RESULT 38
SH1D_CANFA STANDARD; PRT; 377 AA.
AC P11614;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE 5-hydroxytryptamine 1D receptor (5-HT-1D) (Serotonin receptor).
GN HTR1D OR RDC4.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thyroid;
RX MEDLINE=89242119; PubMed=2541503;
RA Libert F., Parmentier M., Lefort A., Dumont J.E., Vassart G.;
RA Maenhaut C., Simons M.-J., Dumont J.E., Vassart G.;
RT "Selective amplification and cloning of four new members of the G
RT protein-coupled receptor family.";
RL Science 244:569-572(1989).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Thyroid;
RX MEDLINE=90245610; PubMed=2159630;
RA Libert F., Parmentier M., Lefort A., Dumont J.E., Vassart G.;
RT "Complete nucleotide sequence of a putative G protein coupled
RT receptor: RDC4.";
RL Nucleic Acids Res. 18:1916-1916(1990).
[3]
RP FUNCTION.
RX MEDLINE=92062181; PubMed=1659418;
RA Maenhaut C., van Sande J., Maesart C., Dinsart C., Libert F.,
RA Monferini E., Giraldo E., Ladinsky H., Vassart G., Dumont J.E.;
RT "The orphan receptor cDNA RDC4 encodes a 5-HT1D serotonin receptor.";
RL Biochem. Biophys. Res. Commun. 180:1460-1468(1991).
[4]
RP FUNCTION.
RX MEDLINE=92100052; PubMed=1758439;
RA Zgombick J.M., Weinshank R.L., Macchi M., Schechter L.E.,
RA Branchek T.A., Hartig P.R.;
RT "Expression and pharmacological characterization of a canine 5-
RT hydroxytryptamine1D receptor subtype.";
RL Mol. Pharmacol. 40:1036-1042(1991).
CC -!- FUNCTION: THIS IS ONE OF THE SEVERAL DIFFERENT RECEPTORS FOR 5-
CC HYDROXYTRYPTAMINE (SEROTONIN), A BIOGENIC HORMONE THAT FUNCTIONS
CC AS A NEUROTRANSMITTER, A HORMONE, AND A MITOGEN. THE ACTIVITY OF
CC THIS RECEPTOR IS MEDIATED BY G PROTEINS THAT INHIBITS ADENYLATE
CC CYCLASE ACTIVITY.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST TO THE OTHER 5HT-1 SUBTYPE RECEPTORS.

```

----- THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch). -----

EMBL; X14049; CAA32207.1; -.
PIR; B30341; B30341.
InterPro; IPR000276; GPCR_Rhodpsn.
Pfam; PF00001; 7tm_1; 1.
PRINTS; PR00237; GPCRHOPOFSN.
PROSITE; PS00237; G-PROTEIN RECP_F1_1; 1.
PROSITE; PS0262; G-PROTEIN RECP_F1_2; 1.
G-protein coupled receptor; Transmembrane; Glycoprotein;
Multigene family.

DOMAIN	1	38	EXTRACELLULAR (POTENTIAL).
FT TRANSMEM	39	62	1 (POTENTIAL).
FT DOMAIN	63	75	CYTOPLASMIC (POTENTIAL).
FT TRANSMEM	76	98	2 (POTENTIAL).
FT DOMAIN	99	108	EXTRACELLULAR (POTENTIAL).
FT TRANSMEM	109	134	3 (POTENTIAL).
FT DOMAIN	135	154	CYTOPLASMIC (POTENTIAL).
FT TRANSMEM	155	176	4 (POTENTIAL).
FT DOMAIN	177	194	EXTRACELLULAR (POTENTIAL).
FT TRANSMEM	195	218	5 (POTENTIAL).
FT DOMAIN	219	302	CYTOPLASMIC (POTENTIAL).
FT TRANSMEM	303	326	6 (POTENTIAL).
FT DOMAIN	327	335	EXTRACELLULAR (POTENTIAL).
FT TRANSMEM	336	360	7 (POTENTIAL).
FT DOMAIN	361	377	CYTOPLASMIC (POTENTIAL).
FT CARBOHYD	5	5	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	17	17	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD	21	21	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID	111	188	BY SIMILARITY.
SQ SEQUENCE	377 AA;	41882 MW;	856406DCE2123EE4 CRC64;

Alignment Scores:
Pred. No.: 20.3 Length: 377
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SH1D_CANFA (1-377)
OY 1917 ATAATCACCAGTGGCCAGCCCTG 1940
Db 47 IleIleThrMetAlaThrAlaLeu 54

RESULT 39
ID -BMP_SOCMV STANDARD; PRT; 462 AA.
AC P15628;
DT 01-APR-1990 (Rel. 14, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Inclusion body matrix protein (Viroplasma).
GN VI.
OS Soybean chlorotic mottle virus.
OC Viruses; Retroviruses; Caulimoviridae;
OC Soybean chlorotic mottle-like viruses.
OX NCBI_TaxID=10651;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90098857; PubMed=2602148;
RA Hasegawa A., Verver J., Shimada A., Saito M., Goldbach R.,
RA van Kammen A., Miki K., Kameya-Iwaki M., Hibi T.;
RT "The complete sequence of soybean chlorotic mottle virus DNA and the
RT identification of a novel promoter.";

CC 2 ferrocyclochrome c.
 CC -1- COFACTOR: CONTAINS TWO PAD AND FOUR TO SIX ZINC MOLES PER MOLE.
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
 CC -1- SIMILARITY: BELONGS TO THE FAD-BINDING OXIDOREDUCTASE/TRANSFERASE
 CC FAMILY 4.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X71628; CAA50635.1; -
 CC PIR; S51528; S51528.
 CC InterPro; IPR004113; FAD-oxidase C.
 CC InterPro; IPR006094; Oxid FAD bind.
 CC Pfam; PF02913; FAD-oxidase C; 1.
 CC Pfam; PF01565; FAD_binding_4; 1.
 CC Oxidoreductase; Flavoprotein; FAD; Transit peptide; Mitochondrion;
 CC Zinc.
 CC -----
 CC TRANSIT 1 ? MITOCHONDRION.
 CC CHAIN ? 579 D-LACTATE DEHYDROGENASE [CYTOCHROME].
 CC SEQUENCE 579 AA; 63484 MW; 0DE3A07DC4934883 CRC64;
 CC -----
 CC Alignment Scores:
 CC Pred. No.: 19.6 Length: 579
 CC Score: 8.00 Matches: 8
 CC Percent Similarity: 100.00% Conservative: 0
 CC Best Local Similarity: 100.00% Mismatches: 0
 CC Query Match: 0.89% Indels: 0
 CC DB: 1 Gaps: 0
 CC -----
 CC US-09-774-490-1 (1-2709) x DLD1_KLJULA (1-579)
 CC QY 495 GAGAGATGATCGAAGTGGCTG 518
 CC DB 470 GluGluMetAsnAlaSerGlyLeu 477
 CC -----
 CC RESULT 42
 CC ID_YE9C SCHPO STANDARD; PRT; 697 AA.
 CC AC 013773;
 CC DT 15-JUL-1998 (Rel. 36, Created)
 CC DT 15-JUL-1998 (Rel. 36, Last sequence update)
 CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
 CC DE Hypothetical J-domain protein C17A5.12 in chromosome 1.
 CC GN SPAC17A5.12.
 CC OS Schizosaccharomyces pombe (Fission yeast).
 CC OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 CC OC Schizosaccharomycetes.
 CC OX NCBI_TaxID=4896;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=972;
 CC RX MEDLINE=21848401; PubMed=11859360;
 CC RA Wood V., Gilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
 CC RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 CC RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 CC RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 CC RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 CC RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 CC RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 CC RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 CC RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitch E.,
 CC RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 CC RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 CC RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 CC RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
 CC RA Welljens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 CC RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,

RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe";
 RL Nature 415:871-880(2002).
 CC -1- SIMILARITY: Contains 1 UBA domain.
 CC -1- CAUTION: THE INITIATOR METHIONINE MAY BE FURTHER DOWNSTREAM.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z98849; CAB11512.1; -
 CC PIR; T37827; T37827.
 CC GenDB SPombe; SPAC17A5.12; -
 CC InterPro; IPR001623; DnaJ_N.
 CC InterPro; IPR001440; TPR.
 CC InterPro; IPR000449; UBA_domain.
 CC Pfam; PF00627; UBA; 1.
 CC SMART; SM00271; DnaJ; 1.
 CC PROSITE; PS00076; DnaJ_2; 1.
 CC PROSITE; PS00030; UBA; 1.
 CC KW Hypothetical protein; Chaperone.
 CC FT DOMAIN 181 221 UBA.
 CC FT DOMAIN 633 696 J-DOMAIN.
 CC SEQUENCE 697 AA; 79483 MW; 43789C80AFC18767 CRC64;
 CC -----
 CC Alignment Scores:
 CC Pred. No.: 19.2 Length: 697
 CC Score: 8.00 Matches: 8
 CC Percent Similarity: 100.00% Conservative: 0
 CC Best Local Similarity: 100.00% Mismatches: 0
 CC Query Match: 0.89% Indels: 0
 CC DB: 1 Gaps: 0
 CC -----
 CC US-09-774-490-1 (1-2709) x YE9C_SCHPO (1-697)
 CC QY 383 TTGCATGAGGAACGGAGTAGGCTG 406
 CC DB 592 LeuAspGluGluArgSerArgLeu 599
 CC -----
 CC RESULT 43
 CC ID_SM4A MOUSE STANDARD; PRT; 760 AA.
 CC AC Q62178;
 CC DT 30-MAY-2000 (Rel. 39, Created)
 CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
 CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
 CC DE Semaphorin 4A precursor (Semaphorin B) (Sema B).
 CC GN SEMA4A OR SEMA4 OR SEMB.
 CC OS Mus musculus (Mouse).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC OX NCBI_TaxID=10090;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=NMRI; TISSUE=Brain;
 CC RX MEDLINE=95267431; PubMed=7748561;
 CC RA Puschel A.W., Adams R.H., Betz H.;
 RT "Marine semaphorin D/collapsin is a member of a diverse gene family
 RT and creates domains inhibitory for axonal extension.";
 RL Neuron 14:941-948(1995).
 CC -1- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO

CC SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- DEVELOPMENTAL STAGE: EXPRESSED FROM DAY 10 IN THE EMBRYO. LOW
 CC LEVELS FOUND BETWEEN DAYS 10-12. EXPRESSION PEAKS ON DAY 13 WITH
 CC MODERATE LEVELS FROM THEN UNTIL BIRTH.
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X85991; CAA59983.1; -;
 CC PIR; I48745; I48745.
 CC MGI; 107560; Sema4a.
 CC InterPro; IPR003659; Plexin-like.
 CC InterPro; IPR002165; Plexin_repeat.
 CC Pfam; PF01437; PSI; 1.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00423; PSI; 1.
 CC SMART; SM00630; Sema; 1.
 CC Signal; Transmembrane; Immunoglobulin domain; Multigene family;
 KW Neurogenesis; Developmental protein; Glycoprotein.
 KW SIGNAL 1 32
 FT CHAIN 33 760
 FT DOMAIN 33 682
 FT TRANSMEM 683 703
 FT DOMAIN 704 760
 FT DOMAIN 64 478
 FT DOMAIN 496 547
 FT DOMAIN 572 630
 FT DOMAIN 609 612
 FT DOMAIN 579 623
 FT DISULFID 120 120
 FT CARBOHYD 135 135
 FT CARBOHYD 496 496
 FT CARBOHYD 606 606
 FT CARBOHYD 606 606
 SQ SEQUENCE 760 AA; 83458 MW; 0A4D6B80767B910F CRC64;
 Alignment Scores:
 Pred. No.: 19.1 Length: 760
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 Db: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x SM4A_MOUSE (1-760)
 QY 578 ACTCACTTGTCGCTGTGGACG 601
 Db 137 ThrHisLeuTyraLcysGlyThr 144
 |||||
 RESULT 44
 SM4A_HUMAN
 ID SM4A_HUMAN STANDARD; PRT; 761 AA.
 AC Q9H3E1; Q9H3E1;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 15-SEP-2003 (Rel. 42, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Semaphorin 4A precursor (Semaphorin B) (Sema B).
 GN Sema4A OR SEMB.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Seki N., Hattori A., Hayashi A., Kozuma S., Muramatsu M., Miyajima N.,
 RA Saito T.;
 RT "Human semaphorin B";
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Colon;
 EX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- FUNCTION: INHIBITS AXONAL EXTENSION BY PROVIDING LOCAL SIGNALS TO
 CC SPECIFY TERRITORIES INACCESSIBLE FOR GROWING AXONS (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
 CC -!- SIMILARITY: Contains 1 Sema domain.
 CC -!- SIMILARITY: Contains 1 immunoglobulin-like C2-type domain.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB029394; BAB20087.1; -;
 CC EMBL; BC020974; AAH20974.1; -;
 CC MIM; 607292; -;
 CC InterPro; IPR003659; Plexin-like.
 CC InterPro; IPR002165; Plexin_repeat.
 CC InterPro; IPR001627; Sema.
 CC Pfam; PF01437; PSI; 1.
 CC Pfam; PF01403; Sema; 1.
 CC SMART; SM00423; PSI; 1.
 CC Signal; Transmembrane; Immunoglobulin domain; Multigene family;
 KW Neurogenesis; Developmental protein; Glycoprotein.
 KW SIGNAL 1 32
 FT CHAIN 33 761
 FT DOMAIN 33 683
 FT TRANSMEM 684 704
 FT DOMAIN 705 761
 FT DOMAIN 64 478
 FT DOMAIN 496 548
 FT DOMAIN 573 631
 FT DISULFID 580 624
 FT CARBOHYD 120 120
 FT CARBOHYD 135 135
 FT CARBOHYD 496 496
 FT CARBOHYD 607 607
 FT CARBOHYD 293 328
 FT CONFLICT 354 354
 FT CONFLICT 354 354
 FT Y -> F (IN REF. 1).
 FT REF. 1)
 FT SAPRSRGSCPTSSATSCSPILLQPLTSTQSSPPSG (IN
 FT CTPQGLPFNVIRHALLPADSPAPRIYAVFTSQW ->


```

DR InterPro; IPR003599; Ig.
DR InterPro; IPR003006; Ig_MHC.
DR InterPro; IPR003659; plexin-like.
DR InterPro; IPR002165; plexin_repeat.
DR InterPro; IPR001627; Sema.
DR Pfam; PF01437; PSI; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; Ig; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KW Signal; Transmembrane; Immunoglobulin domain; Multigene family;
KW Neurogenesis; Developmental protein; Glycoprotein.
FT SIGNAL 1 17
FT CHAIN 18 838
FT DOMAIN 18 675
FT TRANSMEM 676 696
FT DOMAIN 697 838
FT DOMAIN 56 489
FT DOMAIN 507 558
FT DOMAIN 567 649
FT DOMAIN 565 568
FT DOMAIN 763 774
FT DISULFID 584 632
FT CARBOHYD 55 55
FT CARBOHYD 111 111
FT CARBOHYD 126 126
FT CARBOHYD 388 388
FT CARBOHYD 542 542
FT CARBOHYD 598 598
FT CONFLICT 543 543
SQ SEQUENCE 838 AA; 91496 MW; 9B281AEE8681E245 CRC64;

Alignment Scores:
Pred. No.: 18.9 Length: 838
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SM4G_HUMAN (1-838)
QY 578 ACTCACTTGACGCTGTGGAACG 601
DB 128 ThrHisLeuTyrAlaCysGlyThr 135

RESULT 47
FIBI_PETMA
ID FIBI_PETMA STANDARD; PRT; 966 AA.
AC P02674;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Fibrinogen alpha-1 chain precursor [Contains: Fibrinopeptide A]
DE (Fragment).
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
OX NCBI_TaxID=7757;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90122867; PubMed=2611265;
RA Wang Y.Z., Patterson J., Gray J.E., Yu C., Cottrell B.A., Shimizu A.,
RA Graham D., Riley M., Doolittle R.F.;
RT "Complete sequence of the lamprey fibrinogen alpha chain.";
RL Biochemistry 28:9801-9806(1989).
RN [2]
RP SEQUENCE OF 6-11.
RX MEDLINE=77065679; PubMed=998989;
RA Cottrell B.A., Doolittle R.F.;
RT "Amino acid sequences of lamprey fibrinopeptides A and B and
RT characterizations of the junctions split by lamprey and mammalian

```

```

thrombins.";
RL Biochim. Biophys. Acta 453:426-438(1976).
CC -!- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
CC AGGREGATION.
CC -!- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
CC -!- PTM: NOT GLYCOSYLATED.
CC -!- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT. THE SOFT CLOT IS
CC CONVERTED INTO THE HARD CLOT BY FACTOR XIIIa WHICH CATALYZES THE
CC EPSILON-(GAMMA-GLUTAMYL)LYSINE CROSS-LINKING BETWEEN GAMMA CHAINS
CC (STRONGER) AND BETWEEN ALPHA CHAINS (WEAKER) OF DIFFERENT
CC MONOMERS.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.ebi.ac.uk/announcements
or send an email to license@ebi.ac.uk).
-----
EMBL; M30123; AAA49263.1; -.
PIR; A33626; A33626.
PDB; 1LMU; 23-AUG-02.
PDB; 1N73; 07-JAN-03.
KW Blood coagulation; Plasma; Signal; Repeat; 3D-structure.
FT NON TER 1 1
FT SIGNAL <1 5
FT PEPTIDE 6 11
FT CHAIN 12 966
FT SITE 11 12
FIBRINOPEPTIDE A (BY SIMILARITY).
FIBRINOGEN ALPHA-1 CHAIN.
CLEAVAGE (BY THROMBIN; RELEASE
FIBRINOPEPTIDE A) (BY SIMILARITY).
22 X 18 AA APPROXIMATE TANDEM REPEATS.
SER-RICH.
SER-RICH.
SER-RICH.
INTERCHAIN (WITH ALPHA CHAIN)
(BY SIMILARITY).
INTERCHAIN (WITH BETA CHAIN)
(BY SIMILARITY).
INTERCHAIN (WITH GAMMA CHAIN)
(BY SIMILARITY).
INTERCHAIN (WITH BETA CHAIN)
(BY SIMILARITY).
INTERCHAIN (WITH GAMMA CHAIN)
(BY SIMILARITY).
INTERCHAIN (WITH BETA CHAIN)
(BY SIMILARITY).
SQ SEQUENCE 966 AA; 97314 MW; 410520898AA799EE CRC64;

Alignment Scores:
Pred. No.: 18.7 Length: 966
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x FIBI_PETMA (1-966)
QY 153 ACAGGAAGGAGACTAAGCAGCA 176
DB 949 ThrGlyArgLysThrLysAlaAla 956

RESULT 48
SM5A_HUMAN
ID SM5A_HUMAN STANDARD; PRT; 1074 AA.
AC Q13591; O60408;
DT 30-MAY-2000 (Rel. 39, Created)

```


DR Pfam; PF01403; Sema; 1.
 DR Pfam; PF00090; tsp.1; 5.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR SMART; SM00209; TSP1; 6.
 DR PROSITE; PS0092; TSP1; 6.
 DR Signal; Transmembrane; Repeat; Multigene family; Neurogenesis;
 KW Developmental protein; Glycoprotein.
 FT SIGNAL 1 21
 FT CHAIN 22 1077
 FT DOMAIN 22 971
 FT TRANSMEM 972 992
 FT DOMAIN 993 1077
 FT DOMAIN 226 507
 FT DOMAIN 540 593
 FT DOMAIN 595 651
 FT DOMAIN 653 702
 FT DOMAIN 707 765
 FT DOMAIN 784 839
 FT DOMAIN 841 896
 FT DOMAIN 897 944
 FT CARBOHYD 147 147
 FT CARBOHYD 168 168
 FT CARBOHYD 227 227
 FT CARBOHYD 277 277
 FT CARBOHYD 323 323
 FT CARBOHYD 367 367
 FT CARBOHYD 536 536
 FT CARBOHYD 591 591
 FT CARBOHYD 717 717
 FT CARBOHYD 933 933
 SQ SEQUENCE 1077 AA; 120826 MW; 120826 MW; EDABDDDA42789FF CRC64;

Alignment Scores:
 Pred. No.: 19.5 Length: 1077
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SMSA_MOUSE (1-1077)
 QY 929 TACTTTTCTTCGGAATGCA 952
 Db 230 TyrPhePheArgGluAsnAla 237
 |||||

RESULT 50
 RPOC_PSEPU
 ID_RPOC_PSEPU STANDARD; PRT; 1409 AA.
 AC P19176;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (Transcriptase
 DE beta' chain) (RNA polymerase beta' subunit).
 GN RPOC.
 OS Pseudomonas putida.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OC NCBI_TaxID=303;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Danilovich A.V., Borodin A.M., Allikmets R.L., Rostapshov V.M.,
 Chernov I.P., Azhikina T.L., Monastyrskaya G.S., Sverdlov E.D.;
 RT "Nucleotide sequence of the rpoC gene coding for the beta'-subunit of
 RT RNA polymerase in Pseudomonas putida.";
 RL Dokl. Biochem. 303:241-245 (1988).
 RN [2]
 RP SEQUENCE OF 1-497 FROM N.A.
 RX MEDLINE=89117617; PubMed=3219133;
 RA Borodin A.M., Danilovich A.V., Chernov I.P., Azhikina T.L.,
 Rostapshov V.M., Monastyrskaya G.S.;

RT "Genes coding for RNA polymerase in bacteria. III. The use of
 modified Sanger's method for sequencing the C-terminal region of rpoB
 gene, N-terminal region of rpoC gene and intercistron region of rpoA
 polymerase in Pseudomonas putida.";
 RL Bioorg. Khim. 14:1179-1182 (1988).
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 {RNA} (N).
 CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
 ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
 BETA' CHAIN.
 CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 or send an email to license@isb-sib.ch.

 CC EMBL; X16538; CAA34538.1; -;
 CC EMBL; M38319; AAA25987.1; -;
 DR PIR; JN0420; JN0420.
 DR HSPF; Q9KWU6; IHQM.
 DR InterPro; IPR000722; RNA_pol_A.
 DR InterPro; IPR007080; RNA_pol_Rpbi_1.
 DR InterPro; IPR007066; RNA_pol_Rpbi_3.
 DR InterPro; IPR007083; RNA_pol_Rpbi_4.
 DR InterPro; IPR007081; RNA_pol_Rpbi_5.
 DR InterPro; IPR006592; RNA_pol_N.
 DR Pfam; PF04997; RNA_pol_Rpbi_1; 1.
 DR Pfam; PF00623; RNA_pol_Rpbi_2; 1.
 DR Pfam; PF04983; RNA_pol_Rpbi_3; 1.
 DR Pfam; PF05000; RNA_pol_Rpbi_4; 1.
 DR Pfam; PF04998; RNA_pol_Rpbi_5; 1.
 DR SMART; SM00663; RPOLA_N; 1.
 KW Transferase; DNA-directed RNA polymerase; Transcription.
 FT CONFLICT 2 2 P -> L (IN REF. 2).
 FT CONFLICT 17 17 N -> I (IN REF. 2).
 FT CONFLICT 89 89 L -> V (IN REF. 2).
 FT CONFLICT 203 203 I -> T (IN REF. 2).
 FT CONFLICT 328 328 I -> T (IN REF. 2).
 FT CONFLICT 335 335 L -> S (IN REF. 2).
 FT CONFLICT 348 348 L -> R (IN REF. 2).
 FT CONFLICT 482 482 P -> L (IN REF. 2).
 FT CONFLICT 484 484 T -> I (IN REF. 2).
 SQ SEQUENCE 1409 AA; 154544 MW; C5662109D3BE3D81 CRC64;

Alignment Scores:
 Pred. No.: 18.1 Length: 1409
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x RPOC_PSEPU (1-1409)
 QY 2527 TCTAGAACCTCAACACAGTAGAA 2550
 Db 971 SerArgAsnLeuLysGlnValGlu 978
 |||||

RESULT 51
 RYR1_PIG
 ID_RYR1_PIG STANDARD; PRT; 5035 AA.
 AC P16960;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Ryanodine receptor 1 (Skeletal muscle-type ryanodine receptor) (RYR1)

OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Skeletal muscle;
RX MEDLINE=89262082; PubMed=2725677;
RA Takeshima H., Nishimura S., Matsumoto T., Ishido H., Kangawa K.,
RA Minamino N., Matsuo H., Ueda M., Hanaoka M., Hirose T., Numa S.;
RT "Primary structure and expression from complementary DNA of skeletal
RL muscle ryanodine receptor.";
RN Nature 339:439-445(1989).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=90130482; PubMed=2298749;
RA Zorzato F., Fujii J., Otsu K., Phillips M., Green N.M., Lai F.A.,
RA Meissner G., MacLennan D.H.;
RT "Molecular cloning of cDNA encoding human and rabbit forms of the
RT Ca2+ release channel (ryanodine receptor) of skeletal muscle
RT sarcoplasmic reticulum.";
RL J. Biol. Chem. 265:2244-2256(1990).
[3]
RP PHOSPHORYLATION OF SER-2843.
RX MEDLINE=93120178; PubMed=8380342;
RA Suiko J., Maurer-Fogy I., Plank B., Bertel O., Wiskovsky W.,
RA Hohenegger M., Hellmann G.;
RT "Phosphorylation of serine 2843 in ryanodine receptor-calcium release
RT channel of skeletal muscle by cAMP-, cGMP- and CaM-dependent protein
RT kinase.";
RL Biochim. Biophys. Acta 1175:193-206(1993).
CC -!- FUNCTION: COMMUNICATION BETWEEN TRANSVERSE-TUBULES AND
CC SARCOPLASMIC RETICULUM. CONTRACTION OF SKELETAL MUSCLE IS
CC TRIGGERED BY RELEASE OF CALCIUM IONS FROM SR FOLLOWING
CC DEPOLARIZATION OF T-TUBULES.
CC -!- SUBUNIT: Homotetramer (Potential).
CC -!- TISSUE SPECIFICITY: FAST- OR SLOW-TWITCH SKELETAL MUSCLE.
CC -!- PTM: The N-terminus is blocked.
CC -!- MISCELLANEOUS: The calcium release channel is modulated by calcium
CC ions, magnesium ions, ATP and calmodulin.
CC -!- MISCELLANEOUS: The calcium release channel activity resides in the
CC C-terminal region while the remaining part of the protein
CC constitutes the 'foot' structure spanning the junctional gap
CC between the SR and the T-tubule. It is possible that the foot
CC structure interacts with the cytoplasmic region of the
CC dihydropyridine receptor.
CC -!- MISCELLANEOUS: Ryanodine is an alkaloid that binds to the Ca-
CC release channel in junctional SR and modulates its activity.
CC -!- SIMILARITY: BELONGS TO THE RYANODINE RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X15209; CAA33279.1; -;
DR EMBL; X15750; CAA33762.1; -;
DR PIR; S04654; B35041.
DR InterPro; IPR000699; Ca-rel channel.
DR InterPro; IPR001682; Ca/Na_Pore.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR005821; Ion trans.
DR InterPro; IPR003608; MIR.
DR InterPro; IPR001215; Ryanodn_receptor.
DR InterPro; IPR003032; RyR.
DR InterPro; IPR003877; SRP_receptor.
DR Pfam; PF00520; Ion_trans; 1.
DR Pfam; PF02815; MIR; 4.
DR Pfam; PF01365; RYDR_ITPR; 2.

DR Pfam; PF02026; RyR; 4.
DR Pfam; PF00622; SPRY; 3.
DR PRINTS; PR00795; RYANODINER.
DR SMART; SM00472; MIR; 4.
DR SMART; SM00449; SPRY; 3.
KW Receptor; Transmembrane; Ionic channel; Calcium channel; Repeat;
KW Phosphorylation; Glycoprotein.
FT DOMAIN 1 3123
FT TRANSMEM 3124 3144
FT TRANSMEM 3188 3206
FT TRANSMEM 3385 4004
FT TRANSMEM 4023 4041
FT TRANSMEM 4277 4300
FT TRANSMEM 4342 4362
FT TRANSMEM 4559 4580
FT TRANSMEM 4648 4671
FT TRANSMEM 4789 4809
FT TRANSMEM 4837 4856
FT TRANSMEM 4879 4898
FT TRANSMEM 4914 4937
FT DOMAIN 659 797
FT DOMAIN 1085 1208
FT DOMAIN 1430 1570
FT DOMAIN 1873 1913
FT DOMAIN 842 2959
FT REPEAT 842 955
FT REPEAT 956 1069
FT REPEAT 1345 1360
FT REPEAT 1373 1388
FT REPEAT 2726 2845
FT REPEAT 2846 2959
FT MOD_RES 2843 2843
FT MOD_RES 3952 3952
FT MOD_RES 4323 4323
FT CARBOHYD 3466 3466
FT CARBOHYD 3909 3909
FT CARBOHYD 3950 3950
FT CARBOHYD 4149 4149
FT CARBOHYD 4464 4464
FT CONFLICT 2015 2015
FT CONFLICT 3481 3485
SQ SEQUENCE 5037 AA; 565238 MW; 4ABD87AA26697070 CRC64;
Alignment Scores:
Pred. No.: 16.2 Length: 5037
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 1 Gaps: 0
US-09-774-490-1 (1-2709) x RYR1_RABIT (1-5037)
Qy 64 GGGTCGACACCCCTCGACGTGGCA 41
Db 2369 GlySerThrProLeuaspValala 2276
RESULT 53
RYR1 HUMAN
ID RYR1_HUMAN STANDARD; PRT; 5038 AA.
AC P21817; Q16314; Q16368; Q9NPK1; Q9PIU4;
DT 01-MAY-1991 (Rel. 18, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ryanodine receptor 1 (skeletal muscle-type ryanodine receptor) (RyR1
DE (RyR-1) (skeletal muscle calcium release channel)).
GN RYR1 OR RYDR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND PARTIAL SEQUENCE.

RA Adnet P., Haudecoeur G., Krivosic I., McCarthy T.V., Lunardi J.;
 RT "Identification of heterozygous and homozygous individuals with the
 RT novel RYR1 mutation Cys35Arg in a large kindred."
 RL Anesthesiology 86:620-626(1997).
 RN [18]
 RP VARIANT MH LEU-614.
 RX MEDLINE=98051290; PubMed=9389851;
 RA Quane K.A., Ordling H., Keating K.E., Manning B.M., Heine R.,
 RA Bendixen D., Berg K., Krivosic-Horber R., Lehmann-Horn F.,
 RA Fagerlund T.H., McCarthy T.V.;
 RA "Detection of a novel mutation at amino acid position 614 in the
 RT ryanodine receptor in malignant hyperthermia."
 RL Br. J. Anaesth. 79:332-337(1997).
 RN [19]
 RP VARIANT MH TRP-552.
 RX MEDLINE=97284075; PubMed=9138151;
 RA Keating K.E., Giblin L., Lynch P.J., Quane K.A., Lehane M.,
 RA Heffron J.J.A., McCarthy T.V.;
 RA "Detection of a novel mutation in the ryanodine receptor gene in an
 RT Irish malignant hyperthermia pedigree: correlation of the IVCT
 RT response with the affected and unaffected haplotypes."
 RL J. Med. Genet. 34:291-296(1997).
 RN [20]
 RP VARIANTS MH CYS-2163; HIS-2163; MET-2168 AND MET-2206.
 RX MEDLINE=98163444; PubMed=9497245;
 RA Manning B.M., Quane K.A., Ordling H., Urwyler A., Tegazzin V.,
 RA Lehane M., O'Halloran J., Hartung E., Giblin L.M., Lynch P.J.,
 RA Vaughan P., Censier K., Bendixen D., Comi G.P., Heytens L.,
 RA Monseurs K., Fagerlund T.H., Wolz W., Heffron J.J.A., Mueller C.R.,
 RA McCarthy T.V.;
 RA "Identification of novel mutations in the ryanodine-receptor gene
 RT (RYR1) in malignant hyperthermia: genotype-phenotype correlation."
 RL Am. J. Hum. Genet. 62:599-609(1998).
 RN [21]
 RP VARIANTS MH CYS-2459 AND HIS-2459.
 RX MEDLINE=98111378; PubMed=9450902;
 RA Manning B.M., Quane K.A., Lynch P.J., Urwyler A., Tegazzin V.,
 RA Krivosic-Horber R., Censier K., Comi G.P., Adnet P., Wolz W.,
 RA Lunardi J., Muller C.R., McCarthy T.V.;
 RA "Novel mutations at a Cpg dinucleotide in the ryanodine receptor in
 RT Alignment Scores:
 Pred. No.: 16.2 Length: 5038
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x RYR1_HUMAN (1-5038)
 QY 64 CGGTCGACCCCTCGACGTGCA 41
 Db 2269 GlySerThrProLeuAspValala 2276
 RESULT 54
 PSJ1_EUGGR
 ID PSJ1_EUGGR STANDARD; PRT; 37 AA.
 AC P30394;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Photosystem I reaction center subunit IX (PSI-J).
 PSJ1
 OS Euglena gracilis.
 OG Chloroplast.
 OC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.
 OX NCBI TaxID=3039;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Z / UTEX 753;
 RX MEDLINE=89041581; PubMed=3141903;
 RA Manzara T.; Hallick R.B.;

RT "Nucleotide sequence of the Euglena gracilis chloroplast genes for
 RT serine and proline transfer RNAs and a functional open reading
 RL frame";
 RL Nucleic Acids Res. 16:9866-9866(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Z / UTEX 753;
 RX MEDLINE=93169691; PubMed=8435857;
 RA Drager R.G., Hallick R.B.;
 RA "A novel Euglena gracilis chloroplast operon encoding four ATP
 RT synthase subunits and two ribosomal proteins contains 17 introns."
 RL Curr. Genet. 23:271-280(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Z / UTEX 753;
 RX MEDLINE=93347989; PubMed=8346031;
 RA Hallick R.B., Hong L., Drager R.G., Favreau M.R., Monfort A.,
 RA Orsat B., Spielmann A., Stutz E.;
 RA "Complete sequence of Euglena gracilis chloroplast DNA."
 RL Nucleic Acids Res. 21:3537-3544(1993).
 CC 1- FUNCTION: May help in the organization of the psaE and psaF
 CC subunits (By similarity).
 CC -1- SIMILARITY: Belongs to the psaJ family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.ebi-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; Z11874; CAA77934.1; --
 DR EMBL; M18672; AAA84231.1; --
 DR EMBL; X70810; --; NOT_ANNOTATED_CDS.
 DR PIR; S03674; S03674.
 DR HAMAP; MF 00522; --; 1.
 DR InterPro; IPR002615; PSI_Psaj.
 DR Pfam; PF01701; PSI_Psaj; 1.
 DR ProDom; PD004198; PSI_Psaj; 1.
 KW Chloroplast; Photosystem I; Photosynthesis; Transmembrane.
 FT TRANSMEM 7 27
 FT SEQUENCE 37 AA; 4296 MW; 3071BEAA56307EB1 CRC64;
 Alignment Scores:
 Pred. No.: 292 Length: 37
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.78% Indels: 0
 DB: 1 Gaps: 0
 US-09-774-490-1 (1-2709) x PSJ1_EUGGR (1-37)
 QY 725 CTGACGACATCCCTTTTAATA 745
 Db 21 LeuthralaserLeulle 27
 RESULT 55
 RS14_PYRAE
 ID RS14_PYRAE STANDARD; PRT; 54 AA.
 AC O82VWL; P58732;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE 30S ribosomal protein S14P.
 GN RPS14P OR PAE2097.
 OS Pyrobaculum aerophilum.
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
 OC Thermoproteaceae; Pyrobaculum.
 OX NCBI TaxID=13773;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RT Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
aerophilum";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
CC -!- SIMILARITY: BELONGS TO THE S14P FAMILY OF RIBOSOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE009857; AAL63943.1;
DR InterPro; IPR001209; Ribosomal_S14.
DR Pfam; PF00253; Ribosomal_S14; 1.
DR PROSITE; PS00527; RIBOSOMAL_S14; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 54 AA; 6444 MW; 153B7BB801EDD963 CRC64;

Alignment Scores:
Pred. No.: 283 Length: 54
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x RS14_PYRAE (1-54)

Qy 2406 CTTTCGCCACGCTGTTTCGG 2386
Db 36 LeuCysArgArgCysPheArg 42

RESULT 56
YYP2_SYPN2
ID _YYP2_SYPN2 STANDARD; PRT; 63 AA.
AC P32040;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein in pect 3'region (Fragment).
OS Synecococcus sp. (strain PCC 7002) (Agmenellum quadruplicatum).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=32049;
RN [1]
RP SEQUENCE FROM N.A.
RA Schlachter W.M., Bryant D.A.;
RT "Molecular characterization of ferredoxin-NADP+ oxidoreductase in
cyanobacteria: cloning and sequence of the peth gene of Synecococcus
sp. PCC 7002 and studies on the gene product.";
RL Biochemistry 31:3092-3102(1992).
CC -!- SIMILARITY: Contains 1 response regulatory domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M62334; AAA27325.1;
DR InterPro; IPR001789; Response reg.
DR ProDom; PD000039; Response reg; 1.
DR PROSITE; PS00110; RESPONSE_REGULATORY; 1.
KW Hypothetical protein; Sensory transduction.
FT DOMAIN 20 >63 RESPONSE REGULATORY.

FT NON TER 63
SQ SEQUENCE 63 AA; 7142 MW; A9F58D9E01180896 CRC64;

Alignment Scores:
Pred. No.: 279 Length: 63
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x YPH2_SYPN2 (1-63)

Qy 110 GCTGACATGGGAGAACAGGC 90
Db 48 AlaAspAenGlyArgThrGly 54

RESULT 57
YYP3_AZOCA
ID _YYP3_AZOCA STANDARD; PRT; 68 AA.
AC P33988;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in RPO3 3'region (ORF3) (Fragment).
OS Azorhizobium caulinodans.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Hyphomicrobiaceae; Azorhizobium.
OX NCBI_TaxID=7;
RN [1]
RP SEQUENCE FROM N.A.
RA Stigter J., Schneider M., de Bruijn P.J.;
RT "Azorhizobium caulinodans nitrogen fixation (nif/fix) gene
regulation: mutagenesis of the nifA -24/-12 promoter element,
RT characterization of a ntrA(rpoN) gene, and derivation of a model.";
RL Mol. Plant Microbe Interact. 6:238-252(1993).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X69959; CAA49583.1;
DR PIR; S33581; S33581.
KW Hypothetical protein.
FT NON TER 68
SQ SEQUENCE 68 AA; 7612 MW; 435C7F141E9C54EB CRC64;

Alignment Scores:
Pred. No.: 277 Length: 68
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x YRP3_AZOCA (1-68)

Qy 1719 CTGGGGTTCGCCAGCTCCCTT 1739
Db 24 LeuGlyLeuProSerSerLeu 30

RESULT 58
YL13_ARCFU
ID _YL13_ARCFU STANDARD; PRT; 99 AA.
AC O2B167;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)

```
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein AF2113.
GN AF2113.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
CC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Klenk H.-P., Clayton R.A., Tomb J.-F., Hickey E.K., Peterson J.D.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Kyrleides N.C.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrleides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.P., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE000958; AAB89154.1; -
DR PIR; A69514; A69514.
DR TIGR; AF2113; -.
KW Hypothetical protein; Coiled coil; Complete proteome.
FT DOMAIN 43 95 COILED COIL (POTENTIAL).
SQ SEQUENCE 99 AA; 11831 MW; 3E9967F3F16AB65C CRC64;
Alignment Scores:
Pred. No.: 268 Length: 99
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 1 Gaps: 0
US-09-774-490-1 (1-2709) x YL13_ARCFU (1-99)
QY 2210 GAGCATTTGGAGAACTTCTT 2230
DB 6 GluHisLeuGluGluLeu 12
RESULT 59
KF1_COLLI
ID KF1_COLLI STANDARD; PRT; 100 AA.
AC Q9PR15;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Feather keratin Cosi-1/Cosi-3(Cosi-2-1 (F-ker)).
OS Columba livia (Domestic pigeon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Columbiformes; Columbidae; Columba.
OX NCBI_TaxID=8932;
RN (1)
RP SEQUENCE FROM N.A.
RA Takahashi R.;
RT "Pigeon feather keratin genes.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: THE AVIAN KERATINS (F-KER, S-KER, C-KER AND B-KER) ARE A
```

```
CC COMPLEX MIXTURE OF VERY SIMILAR POLYPEPTIDES.
CC -!- SIMILARITY: BELONGS TO THE AVIAN KERATIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB017901; BAA33467.1; -
DR EMBL; AB017903; BAA33469.1; -
DR EMBL; AB017904; BAA33470.1; -
DR InterPro; IPR003461; Keratin.
DR Pfam; PF02422; Keratin; 1.
KW Keratin; Multigene family; Feather; Acetylation.
FT INIT MET 0 BY SIMILARITY.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
SQ SEQUENCE 100 AA; 10235 MW; 54292E196A53DF99 CRC64;
Alignment Scores:
Pred. No.: 268 Length: 100
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 1 Gaps: 0
US-09-774-490-1 (1-2709) x KRF1_COLLI (1-100)
QY 2512 CTGAGCTGCATTACTCTAGA 2532
DB 85 LeuSerCysIleThrSerArg 91
RESULT 60
KF2_COLLI
ID KF2_COLLI STANDARD; PRT; 100 AA.
AC Q93439;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Feather keratin Cosi-2 (F-ker).
OS Columba livia (Domestic pigeon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Columbiformes; Columbidae; Columba.
OX NCBI_TaxID=8932;
RN (1)
RP SEQUENCE FROM N.A.
RA Takahashi R.;
RT "Pigeon feather keratin genes.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SUBUNIT: THE AVIAN KERATINS (F-KER, S-KER, C-KER AND B-KER) ARE A
CC COMPLEX MIXTURE OF VERY SIMILAR POLYPEPTIDES.
CC -!- SIMILARITY: BELONGS TO THE AVIAN KERATIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB017902; BAA33468.1; -
DR InterPro; IPR003461; Keratin.
DR Pfam; PF02422; Keratin; 1.
KW Keratin; Multigene family; Feather; Acetylation.
FT INIT MET 0 BY SIMILARITY.
FT MOD_RES 1 1 ACETYLATION (BY SIMILARITY).
SQ SEQUENCE 100 AA; 10166 MW; 53C50756254E5CED CRC64;
Alignment Scores:
```

Pred. No.: 268 Length: 100
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x KRF2_COLLI (1-100)

QY 2512 CTGAGCTGCATTACTCTAGA 2532

DB 85 LeuSerCysIleThrSerArg 91

RESULT 61

KRF3_COLLI

ID KRF3_COLLI STANDARD; PRT; 100 AA.

AC O93500;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Feather keratin Cos2-2 (F-ker).

OS Columba livia (Domestic pigeon).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Columbiformes; Columba.

OX NCBI_TaxID=8932;

RN [1]

RP SEQUENCE FROM N.A.

RA Takahashi R.;

RT "Pigeon feather keratin genes.";

RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

CC -!- SUBUNIT: THE AVIAN KERATINS (F-KER, S-KER, C-KER AND B-KER) ARE A

CC COMPLEX MIXTURE OF VERY SIMILAR POLYPEPTIDES.

CC -!- SIMILARITY: BELONGS TO THE AVIAN KERATIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AB017905; BAA33471.1; -.

DR InterPro; IPR003461; Keratin.

DR Pfam; PF02422; Keratin; 1.

KW Keratin; Multigene family; Feather; Acetylation.

FT INIT MET 0 BY SIMILARITY

FT MOD_RES 1 ACETYLATION (BY SIMILARITY).

SQ SEQUENCE 100 AA; 10197 MW; AS8E713B92D9E065 CRC64;

Alignment Scores:

Pred. No.: 268 Length: 100
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x KRF3_COLLI (1-100)

QY 2512 CTGAGCTGCATTACTCTAGA 2532

DB 85 LeuSerCysIleThrSerArg 91

RESULT 62

KRF4_COLLI

ID KRF4_COLLI STANDARD; PRT; 100 AA.

AC Q9PSV3;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Feather keratin Cos2-3 (F-ker).

OS Columba livia (Domestic pigeon).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Columbiformes; Columba.
OX NCBI_TaxID=8932;

RN [1]

RP SEQUENCE FROM N.A.

RA Takahashi R.;

RT "Pigeon feather keratin genes.";

RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.

CC -!- SUBUNIT: THE AVIAN KERATINS (F-KER, S-KER, C-KER AND B-KER) ARE A

CC COMPLEX MIXTURE OF VERY SIMILAR POLYPEPTIDES.

CC -!- SIMILARITY: BELONGS TO THE AVIAN KERATIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; AB017906; BAA33472.1; -.

DR InterPro; IPR003461; Keratin.

DR Pfam; PF02422; Keratin; 1.

KW Keratin; Multigene family; Feather; Acetylation.

FT INIT MET 0 BY SIMILARITY

FT MOD_RES 1 ACETYLATION (BY SIMILARITY).

SQ SEQUENCE 100 AA; 10288 MW; BB885DF31EE0952F CRC64;

Alignment Scores:

Pred. No.: 268 Length: 100
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x KRF4_COLLI (1-100)

QY 2512 CTGAGCTGCATTACTCTAGA 2532

DB 85 LeuSerCysIleThrSerArg 91

RESULT 63

YC54_AQUAE

ID YC54_AQUAE STANDARD; PRT; 102 AA.

AC O67255;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein AQ_1254.

GN AQ_1254.

OS Aquifex aeolicus.

OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.

OX NCBI_TaxID=63363;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=VF5;

RX MEDLINE=98196666; PubMed=9537320;

RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,

RA Graham D.E., Overbeek R., Sneed M.A., Keller M., AuJay M., Huber R.,

RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;

RT "The complete genome of the hyperthermophilic bacterium Aquifex

aeolicus.";

RL Nature 392:353-358 (1998).

CC -!- SIMILARITY: BELONGS TO THE UPF0092 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

```

CC -----
DR EMBL: AE000731; AAC07266.1; -
DR PIR: D70408; D70408; YajC.
DR InterPro: IPR003849; YajC.
DR Pfam: PF02699; YajC; 1.
DR TIGRFAMs: TIGR00739; YajC; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 9 29 POTENTIAL.
SQ SEQUENCE 102 AA; 11823 MW; 2361BFC80B2F3277 CRC64;

Alignment Scores:
Pred. No.: 268 Length: 102
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x YC54_AQUAE (1-102)
QY 1404 AGGTTATACATCATCAGGA 1384
DB 51 LysValIleThrSerSerGly 57

RESULT 64
INS2_XENLA
ID INS2_XENLA STANDARD; PRT; 106 AA.
AC P12707;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Insulin 2 precursor.
GN INS2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89255444; PubMed=2722842;
RA Shuldiner A.R., Phillips S., Roberts C.T. Jr., Lerioth D., Roth J.;
RT "Xenopus laevis contains two nonallelic preproinsulin genes. cDNA
RT cloning and evolutionary perspective.";
RL J. Biol. Chem. 264:9428-9432(1989).
RN [2]
RP SEQUENCE OF 24-53 AND 86-106.
RX MEDLINE=89289601; PubMed=2661211;
RA Shuldiner A.R., Bennett C., Robinson E.A., Roth J.;
RT "Isolation and characterization of two different insulins from an
RT amphibian, Xenopus laevis.";
RL Endocrinology 125:469-477(1989).
CC -!- FUNCTION: INSULIN DECREASES BLOOD GLUCOSE CONCENTRATION. IT
CC INCREASES CELL PERMEABILITY TO MONOSACCHARIDES, AMINO ACIDS AND
CC FATTY ACIDS. IT ACCELERATES GLYCOLYSIS, THE PENTOSE PHOSPHATE
CC CYCLE, AND GLYCOGEN SYNTHESIS IN LIVER.
CC -!- SUBUNIT: HETERODIMER OF A B CHAIN AND AN A CHAIN LINKED BY TWO
CC DISULFIDE BONDS.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: BELONGS TO THE INSULIN/IGF/RELAXIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M24442; AAA49887.1; -
DR PIR: B33847; IPXL2.
DR HSSP: P01315; 1SDB.
DR InterPro: IPR004825; Ins/IGF/relax.

```

```

DR Pfam: PF00049; Insulin; 1.
DR SMART: SM00078; ILGF; 1.
DR PROSITE: PS00262; INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism; Multigene family; Signal.
FT SIGNAL 1 23
FT CHAIN 24 53 INSULIN 2 B CHAIN.
FT PROPEP 56 83 INSULIN 2 C PEPTIDE.
FT CHAIN 86 106 INSULIN 2 A CHAIN.
FT DISULFID 30 92 INTERCHAIN.
FT DISULFID 42 105 INTERCHAIN.
FT DISULFID 91 96
SQ SEQUENCE 106 AA; 12207 MW; 15859A0C758E7DEF CRC64;

Alignment Scores:
Pred. No.: 267 Length: 106
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x INS2_XENLA (1-106)
QY 1977 GTGCTACTATTCTCTACACCA 1957
DB 13 ValLeuLeuPheSerThrPro 19

RESULT 65
Y094_HAEIN
ID Y094_HAEIN STANDARD; PRT; 106 AA.
AC P43939;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein HI0094.
GN HI0094.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Rd / KW20 / ATCC 51907;
MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32695; AAC21779.1; -
DR PIR: B64001; B64001.
DR TIGR: HI0094; -
DR InterPro: IPR003474; Glcn transporter.
DR Pfam: PF02447; GntP_permease; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 78 98 POTENTIAL.
SQ SEQUENCE 106 AA; 11735 MW; 672EC13E1F59E3E4 CRC64;

```

Alignment Scores:
 Pred. No.: 267
 Score: 7.00
 Length: 106
 Matches: 7
 Percent Similarity: 100.00%
 Conservative: 0
 Best Local Similarity: 100.00%
 Mismatches: 0
 Query Match: 0.80%
 Indels: 0
 DB: 1
 Gaps: 0

US-09-774-490-1 (1-2709) x Y094_HAEIN (1-106)

QY 283 ATTGTTCTCCATCTGATA 263

Db 89 llevalleuProileLeulle 95

RESULT 66

LHG2_RHOCA
 ID_LHG2_RHOCA STANDARD; PRT; 118 AA.

AC P23460;

DT 01-NOV-1991 (Rel. 20, Created)

DT 01-NOV-1991 (Rel. 20, Last sequence update)

DT 01-NOV-1991 (Rel. 20, Last annotation update)

DE Light-harvesting protein B-800/850, gamma chain.

GN PUCR.

OS Rhodobacter capsulatus (Rhodopseudomonas capsulata).

OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;

OC Rhodobacteraceae; Rhodobacter.

OX NCBI_TaxID=1061;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=89359127; PubMed=2549005;

RA Tichy H.V., Oberle B., Stiehle H., Schiltz E., Drews G.;

RT "Genes downstream from pucB and pucA are essential for formation of

the B800-850 complex of Rhodobacter capsulatus.";

RL J. Bacteriol. 171:4914-4922(1989)

CC -!- FUNCTION: SEEMS TO BE REQUIRED FOR THE LH-II STABILIZATION.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC -----

DR EMBL; M28510; AAA26165.1; -.

DR PIR; E33958; E33958.

KW Antenna complex; Light-harvesting polypeptide.

SQ SEQUENCE 118 AA; 13493 MW; 749AF4CC1EF82262 CRC64;

Alignment Scores:

Pred. No.: 264
 Score: 7.00
 Length: 118
 Matches: 7
 Percent Similarity: 100.00%
 Conservative: 0
 Best Local Similarity: 100.00%
 Mismatches: 0
 Query Match: 0.80%
 Indels: 0
 DB: 1
 Gaps: 0

US-09-774-490-1 (1-2709) x LHG2_RHOCA (1-118)

QY 1685 TGGAAAGTCCTGCTGATA 1665

Db 84 TrpLysAlaProLeuLeuLys 90

RESULT 67

CRCB_NEIMA

ID_CRCB_NEIMA STANDARD; PRT; 119 AA.

AC Q9JUL1;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Protein crCB homolog.

GN CRCB OR NNA1264.

OS Neisseria meningitidis (serogroup A).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;
 OC Neisseriaceae; Neisseria.

OX NCBI_TaxID=65699;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Z2491 / Serogroup A / Serotype 4A;

RX MEDLINE=20222556; PubMed=10761919;

RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,

RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,

RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,

RA Jagels K., Leather S., Moule S., Mungall K., Quail M.A.,

RA Rajandream M.A., Rutherford K.M., Simmonds M., Skelton J.,

RA Whitehead S., Spratt B.G., Barrell B.G.;

RT "Complete DNA sequence of a serogroup A strain of Neisseria

meningitidis Z2491.";

RL Nature 404:502-506(2000).

CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -!- SIMILARITY: BELONGS TO THE CRCB FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC -----

DR EMBL; AL162755; CAB84519.1; -.

DR PIR; F81894; F81894.

DR HAWAP; MF 00454; -; 1.

DR InterPro; IPR003691; Camphor_Crcb.

DR Pfam; PF02537; CRCB; 1.

KW Transmembrane; Complete proteome.

FT TRANSMEM 5 22 POTENTIAL.

FT TRANSMEM 37 54 POTENTIAL.

FT TRANSMEM 61 83 POTENTIAL.

FT TRANSMEM 93 112 POTENTIAL.

SQ SEQUENCE 119 AA; 12456 MW; D3104F0B06F5BD77 CRC64;

Alignment Scores:

Pred. No.: 264
 Score: 7.00
 Length: 119
 Matches: 7
 Percent Similarity: 100.00%
 Conservative: 0
 Best Local Similarity: 100.00%
 Mismatches: 0
 Query Match: 0.80%
 Indels: 0
 DB: 1
 Gaps: 0

US-09-774-490-1 (1-2709) x CRCB_NEIMA (1-119)

QY 1167 ACTGGATTTTGGATCTTTA 1147

Db 65 ThrGlyPheLeuGlySerLeu 71

RESULT 68

CRCB_NEIMB

ID_CRCB_NEIMB STANDARD; PRT; 119 AA.

AC Q9JZG6;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Protein crCB homolog.

GN CRCB OR NMB1065.

OS Neisseria meningitidis (serogroup B).

OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;

OC Neisseriaceae; Neisseria.

OX NCBI_TaxID=491;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=MC58 / Serogroup B;

RX MEDLINE=20175755; PubMed=10710307;

RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,

RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,

RA Nelson W.C., Gwinn M.L., DeBoy R., Peterson J.D., Hickey E.K.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Citron H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scariato V., Maignani V., Pizzo M., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RT "Complete genome sequence of *Neisseria meningitidis* serogroup B strain
MC58.";
RL Science 287:1809-1815(2000).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE CRCB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE002457; AAF41461.1; -;
DR PIR; C81125; C81125.
DR TIGR; NMB1065; -;
DR HAMAP; MF_00454; -; 1.
DR InterPro; IPR003691; Camphor_CrcB.
DR Pfam; PF02537; CRCB; 1.
KW Transmembrane; Complete proteome.
FT TRANSMEM 5 22 POTENTIAL.
FT TRANSMEM 37 54 POTENTIAL.
FT TRANSMEM 61 83 POTENTIAL.
FT TRANSMEM 93 112 POTENTIAL.
SQ SEQUENCE 119 AA; 12474 MW; D651C88215B9B80B CRC64;

Alignment Scores:
Pred. No.: 264 Length: 119
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x CRCB_NEIMB (1-119)
QY 1167 ACTGGATTTTAGGATCTTTA 1147
|||||
Db 65 ThrGlyPheLeuGlySerLeu 71

RESULT 69
CRCB_PASMU
ID _CRCB_PASMU STANDARD; PRT; 128 AA.
AC Q9CKI7.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Protein crCB homolog.
GN CRCB OR PM1630.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC Pasteurellaceae; Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Pm70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
RA "Complete genomic sequence of *Pasteurella multocida* Pm70.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE CRCB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AE006200; AAK03714.1; ALT_INIT.
DR HAMAP; MF_00454; -; 1.
DR InterPro; IPR003691; Camphor_CrcB.
DR Pfam; PF02537; CRCB; 1.
DR TIGR; TIGR00494; CRCB; 1.
KW Transmembrane; Complete proteome.
FT TRANSMEM 7 26 POTENTIAL.
FT TRANSMEM 36 58 POTENTIAL.
FT TRANSMEM 70 92 POTENTIAL.
FT TRANSMEM 102 124 POTENTIAL.
SQ SEQUENCE 128 AA; 14139 MW; 13D3280EA5A1E805 CRC64;

Alignment Scores:
Pred. No.: 262 Length: 128
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x CRCB_PASMU (1-128)
QY 1167 ACTGGATTTTAGGATCTTTA 1147
|||||
Db 74 ThrGlyPheLeuGlySerLeu 80

RESULT 70
ATP6_THUOB
ID _ATP6_THUOB STANDARD; PRT; 133 AA.
AC Q36090;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE ATP synthase A chain (EC 3.6.3.14) (Protein 6) (Fragment).
GN MTATP6 OR ATP6.
OS *Thunnus obesus* (Bigeye tuna).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC Scombridae; Thunnus.
OX NCBI_TaxID=8241;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96139007; PubMed=8587119;
RA Chow S., Kishino H.;
RA "Phylogenetic relationships between tuna species of the genus *Thunnus*
RA (Scombridae: Teleostei): inconsistent implications from morphology,
RA nuclear and mitochondrial genomes.";
RL J. Mol. Evol. 41:741-748(1995).
CC -1- FUNCTION: KEY COMPONENT OF THE PROTON CHANNEL; IT MAY PLAY A
CC DIRECT ROLE IN THE TRANSLOCATION OF PROTONS ACROSS THE MEMBRANE.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
CC H(+) (Out).
CC -1- SUBUNIT: F-TYPE ATPASES HAVE 2 COMPONENTS, CF(1) - THE CATALYTIC
CC CORE - AND CF(0) - THE MEMBRANE PROTON CHANNEL. CF(1) HAS FIVE
CC SUBUNITS: ALPHA(3), BETA(3), GAMMA(1), DELTA(1), EPSILON(1). CF(0)
CC HAS THREE MAIN SUBUNITS: A, B AND C.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO THE ATPASE A CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----

```
-----
CC EMBL; D63416; BAA09721.1; -
DR InterPro; IPR000568; ATP-Synt Asub.
DR Pfam; PF00119; ATP-synt A; 1.
DR PRINTS; PR00123; ATPASEA.
DR PROSITE; PS00449; ATPASE_A; PARTIAL.
KW Hydrogen ion transport; CF(0); Mitochondrion; Transmembrane.
FT NON TER 1
FT NON TER 133
SQ SEQUENCE 133 AA; 14982 MW; D625359D2374084B CRC64;

Alignment Scores:
Pred. No.: 262 Length: 133
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x ATP6_THUOB (1-133)
OY 1248 CTTCTCATCATCATCATGCTA 1228
Db 51 LeuLeuThrSerLeuMetLeu 57

RESULT 71
SRP_CHLPS
ID - SRP_CHLPS STANDARD; PRT; 134 AA.
AC P28164;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Sulfur-rich protein.
GN SRP.
OS Chlamydia psittaci (Chlamydophila psittaci).
OC Bacteriia; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydophila.
OX NCBI_TaxID=833554;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=6BC;
RX MEDLINE=91267949; PubMed=2050637;
RA Everett K.D.E., Hatch T.P.;
RT "Sequence analysis and lipid modification of the cysteine-rich
envelope proteins of Chlamydia psittaci 6BC.";
RL J. Bacteriol. 173:3821-3830(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M61116; AAB61620.1; -
SQ SEQUENCE 134 AA; 14732 MW; 2982CGFC51BF854 CRC64;

Alignment Scores:
Pred. No.: 261 Length: 134
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x SRP_CHLPS (1-134)
OY 592 GCGGTACAGTGCAGTCTGATT 572
Db 29 GlyValGlnValSerLeuIle 35

RESULT 72
YBJ5_CANAL
```

```
YBJ5_CANAL STANDARD; PRT; 134 AA.
ID PS3716;
AC 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein in WHS11 5'region (Fragment).
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN 1
RP SEQUENCE FROM N.A.
RC STRAIN=WO-1;
RX MEDLINE=93380673; PubMed=7916716;
RA Srikantha T., Soll D.R.;
RT "A white-specific gene in the white-opaque switching system of
Candida albicans.";
RL Gene 131:53-60(1993).
RN 2
RP CONCEPTUAL TRANSLATION.
RA Bairoch A.;
RL Unpublished observations (AUG-1996).
CC 1- SIMILARITY: STRONG, TO YEAST YBL0835.
CC 1- CAUTION: THIS IS A CONCEPTUAL TRANSLATION; FRAMESHIFTS WERE
INTRODUCED TO MAXIMIZE THE SIMILARITY WITH YEAST YBL095W.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S65451; -; NOT_ANNOTATED_CDS.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 134 AA; 14778 MW; 58AE66F982E866FA CRC64;

Alignment Scores:
Pred. No.: 261 Length: 134
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x YBJ5_CANAL (1-134)
OY 728 ACAGCATCCCTTTTAAATAGAT 748
Db 62 ThrAlaSerLeuLeuIleAsp 68

RESULT 73
UGR1_MOUSE
ID UGR1_MOUSE STANDARD; PRT; 139 AA.
AC Q920H1; Q920H2; Q920H3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uteroglobin-related protein 1 precursor (Secretoglobulin family 3A
member 2).
GN SCGB3A2 OR UGRP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 1
RP SEQUENCE FROM N.A. (ISOFORMS A; B AND C).
RC TISSUE=Lung;
RX MEDLINE=21539178; PubMed=11682631;
RA Niimi T., Keck-Waggoner C.L., Popescu N.C., Zhou Y., Levitt R.C.,
RA Kimura S.;
```

RT "UGRP1, a uteroglobin/clara cell secretory protein-related protein, is
 RT a novel lung-enriched downstream target gene for the T/EBP/NKX2.1
 RT homeodomain transcription factor.";
 RL Mol. Endocrinol. 15:2021-2036(2001).
 CC -!- SUBUNIT: Homodimer.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Name=C;
 CC IsoId=Q920H1-1; Sequences=Displayed;
 CC Name=A;
 CC IsoId=Q920H1-2; Sequences=VSP_006727, VSP_006728;
 CC Name=B;
 CC IsoId=Q920H1-3; Sequences=VSP_006726;
 CC -!- TISSUE SPECIFICITY: Highly expressed in lung.
 CC -!- SIMILARITY: BELONGS TO THE UTEROGLOBIN FAMILY. UGRP SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF274959; AAL25708.1; -;
 CC EMBL; AF274960; AAL25709.1; -;
 CC EMBL; AF274961; AAL25710.1; -;
 CC MGD; MGI:2153470; Scg93a2.
 CC GO; GO:0005576; C:extracellular; IDA.
 CC GO; GO:0005515; F:protein binding activity; IPI.
 CC InterPro; IPR006038; Uteroglobin_supf.
 CC Pfam; PF01099; Uteroglobin; 1.
 CC Signal; Alternative splicing.
 CC FT SIGNAL 1 21
 CC FT CHAIN 22 139 UTEROGLOBIN-RELATED PROTEIN 1.
 CC FT VARSPLIC 107 139 VSVLFLPMICAYPRDSKQTFATIERVFEQSKL-> EALS
 CC
 CC FT VARSPLIC 85 91 /FTID=VSP_006726.
 CC FT VARSPLIC 92 139 /FTID=VSP_006727.
 CC FT VARSPLIC 92 139 Missing (in isoform A).
 CC FT /FTID=VSP_006728.
 CC SQ SEQUENCE 139 AA; 15431 MW; 8A3FB080B41E65E4 CRC64;
 CC
 CC Alignment Scores:
 CC Pred. No.: 261 Length: 139
 CC Score: 7.00 Matches: 7
 CC Percent Similarity: 100.00% Conservative: 0
 CC Best Local Similarity: 100.00% Mismatches: 0
 CC Query Match: 0.80% Indels: 0
 CC DB: 1 Gaps: 0
 CC
 CC US-09-774-490-1 (1-2709) x UGR1_MOUSE (1-139)
 CC
 CC QY 1212 GATCCCTTGAAATGCTACTG 1192
 CC |||||
 CC Db 45 AspProLeuLysMetLeuLeu 51
 CC
 CC RESULT 74
 CC HBA5_XENLA
 CC ID_HBA5_XENLA STANDARD; PRT; 141 AA.
 CC AC P06638;
 CC DT 01-JAN-1988 (Rel. 06, Created)
 CC DT 01-APR-1988 (Rel. 07, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Hemoglobin alpha-5 chain (Alpha-TS).
 CC OS Xenopus laevis (African clawed frog).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 CC OC Xenopodinae; Xenopus.
 CC OX NCBI_TaxID=8355;
 CC RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=85297748; PubMed=2993998;
 RA Barville D.; Williams J.G.;
 RT "The pattern of expression of the Xenopus laevis tadpole alpha-globin
 RT genes and the amino acid sequence of the three major tadpole alpha-
 RT globin polypeptides.";
 RL Nucleic Acids Res. 13:5407-5421(1985).
 CC -!- FUNCTION: THIS IS A LARVAL (TADPOLE) ALPHA-GLOBIN.
 CC -!- SUBUNIT: Heterotetramer of two alpha chains and two beta chains.
 CC -!- TISSUE SPECIFICITY: Red blood cells.
 CC -!- SIMILARITY: Belongs to the globin family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X02798; CAA26566.1; -;
 CC PIR; C24338; C24338.
 CC HSSP; P01922; LABW.
 CC InterPro; IPR002338; Alpha haem.
 CC InterPro; IPR000971; Globin.
 CC Pfam; PF00042; globin; 1.
 CC PRINTS; PR00612; ALPHAHAE.
 CC PROSITE; PS01033; GLOBIN; 1.
 CC KW Heme; Oxygen transport; Transport; Erythrocyte.
 CC FT INIT_MET 0 0
 CC FT METAL 58 58 IRON (HEME DISTAL LIGAND).
 CC FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
 CC SQ SEQUENCE 141 AA; 15270 MW; 73E02DE6BF6CEB CRC64;
 CC
 CC Alignment Scores:
 CC Pred. No.: 260 Length: 141
 CC Score: 7.00 Matches: 7
 CC Percent Similarity: 100.00% Conservative: 0
 CC Best Local Similarity: 100.00% Mismatches: 0
 CC Query Match: 0.80% Indels: 0
 CC DB: 1 Gaps: 0
 CC
 CC US-09-774-490-1 (1-2709) x HBA5_XENLA (1-141)
 CC
 CC QY 1346 TGGCACAAGTTCCTGGCCGTG 1326
 CC |||||
 CC Db 125 TrpAspLysPheLeuAlaVal 131
 CC
 CC RESULT 75
 CC PSG8_HUMAN
 CC ID_PSG8_HUMAN STANDARD; PRT; 144 AA.
 CC AC Q9UQ74;
 CC DT 30-MAY-2000 (Rel. 39, Created)
 CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
 CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
 CC DE Pregnancy-specific beta-1-glycoprotein 8 precursor (PSBG-8)
 CC DE (Fragment).
 CC GN PSBG.
 CC OS Homo sapiens (Human).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CC OX NCBI_TaxID=9606;
 CC RN [1]
 CC SEQUENCE FROM N.A.
 CC RA Frangemyr L.; Teglund S.; Israelsson A.; Hammarstrom S.;
 CC RT "Characterization of upstream promoter region, exon1 and exon2 of the
 CC PSB gene family.";
 CC RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: Secreted (Potential).
 CC -!- DEVELOPMENTAL STAGE: PSBG ARE PRODUCED IN HIGH QUANTITY DURING
 CC PREGNANCY.
 CC -!- SIMILARITY: BELONGS TO THE IMMUNOGLOBULIN SUPERFAMILY. CEA
 CC SUBFAMILY.


```

CC  -!- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC  -!- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AF106556; AAD21022.1; -
DR  EMBL; AF106555; AAD21022.1; JOINED.
DR  Genew; HGNC:9525; PSG8.
DR  MIM; 176397; -
DR  InterPro; IPR003599; IG.
DR  SMART; SM00409; IG; 1.
DR  PROSITE; PS50835; IG_LIKE; FALSE NEG.
KW  Immunoglobulin domain; Glycoprotein; Signal; Repeat; Multigene family.
FT  SIGNAL 1 34 POTENTIAL.
FT  CHAIN 35 >144 PREGNANCY-SPECIFIC BETA-1-GLYCOPROTEIN 8.
FT  DOMAIN 35 144 IG-LIKE V-TYPE.
FT  CARBOHYD 61 61 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 104 104 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  CARBOHYD 111 111 N-LINKED (GLCNAC. .) (POTENTIAL).
FT  NON_TER 144 144
SQ  SEQUENCE 144 AA; 16109 MW; 0A83A6711042C1FE CRC64;

Alignment Scores:
Pred. No.: 260 Length: 144
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PSG8_HUMAN (1-144)
Qy 722 CTGCTGACAGCATCCCTTTTA 742
Db 19 LeuLeuThrAlaSerLeuLeu 25

```

Search completed: August 3, 2003, 10:32:23
 Job time : 127 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 10:29:37 ; Search time 128.5 Seconds

(without alignments)
10880.381 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 895

Sequence: 1 aattttttttttatcatgatg.....agggttttttttcttaataacc 2709

Scoring table:

OLIGO Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 830525 seqs, 258052604 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1660916

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters: -MODEL=frame+n2p.model -DEV=xlh
-Q=/csn2.1/USPTO/spool/US09774490/runat_03082003_102927_4459/app_query.fasta_1.2887
-DB=SPTRMBL_23 -QWTF=fastan -SUFFIX=oligo.rspt -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=100 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=75 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HSPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09774490 @CCN 1.1.0 @runat_03082003_102927_4459 -NCPU=6 -ICPU=3
-NO.MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSEBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

SPTRMBL_23:*
1: sp_alchea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mbc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	43	4.8	774	13 Q8JIW9	Q8JIW9 xenopus lae

Q9ns98	homo sapien	782	4	Q9NS98
Q9hbr1	homo sapien	457	4	Q9HBR1
Q96gx0	homo sapien	635	4	Q96GX0
Q8tb71	homo sapien	748	4	Q8TB71
Q8tdv7	homo sapien	754	4	Q8TDV7
Q8bkq6	mus musculus	403	11	Q8BKQ6
Q8bm6	mus musculus	777	11	Q8BMF6
Q8bh34	mus musculus	777	11	Q8BH34
Q9yhx3	brachydanio	328	13	Q9YHX3
Q8ncq1	homo sapien	416	4	Q8NCO1
Q8qgu9	gallus gall	756	13	Q8QGU9
Q9qz23	mus musculus	775	11	Q9QZ23
Q8nas4	homo sapien	132	4	Q8NAS4
Q9ji29	rattus norv	296	11	Q9JI29
Q8axa8	gallus gall	160	13	Q8AXA8
Q64906	alcelaphine	653	12	Q64906
Q9c0c4	homo sapien	963	4	Q9C0C4
Q96qn2	homo sapien	139	4	Q96QN2
Q49704	mycobacteri	223	2	Q49704
Q54948	mus musculus	284	11	Q54948
Q33142	mycobacteri	311	16	Q33142
Q9bw98	homo sapien	325	4	Q9BW98
Q8bn9	homo sapien	333	4	Q8BN9
Q8f7c6	leptospira	350	16	Q8F7R6
Q8wvp9	homo sapien	458	4	Q8WVP9
Q96fk5	homo sapien	475	4	Q96FK5
Q9nx92	homo sapien	510	4	Q9NX92
Q8bi3	mus musculus	550	11	Q8BIC3
Q8bjc1	mus musculus	799	11	Q8BJC1
Q8bif6	mus musculus	823	11	Q8BIR6
Q9c0b8	homo sapien	893	4	Q9C0B8
Q8r3m7	mus musculus	911	11	Q8R3M7
Q8r1b4	mus musculus	911	11	Q8R1B4
Q95201	mustela vis	4859	6	Q95201
Q91319	rana catesb	4868	13	Q91319
Q909b5	gallus gall	4869	13	Q909B5
Q9t633	oryctolagus	4872	6	Q9TS33
Q9ern6	mus musculus	4967	11	Q9ERN6
Q29621	oryctolagus	4968	6	Q29621
Q91313	rana catesb	5037	13	Q91313
Q13054	makaira nig	5081	13	Q13054
Q24498	drosophila	5107	5	Q24498
Q8mks5	drosophila	5113	5	Q8MK53
Q8mks4	drosophila	5127	5	Q8MK54
Q8k4m4	mus musculus	48	11	Q8K4M4
Q924c2	shigella so	51	2	Q924C2
Q9335	pyrococcus	101	17	Q9335
Q8ze0	versinia pe	178	16	Q8ZE0
Q9s7s7	arabidopsis	180	10	Q9S7S7
Q9mb00	brassica ca	184	10	Q9MB00
Q8cl36	versinia pe	218	16	Q8CL36
Q9kam9	bacillus ha	223	16	Q9KAM9
Q8hm61	lophius ame	227	8	Q8HM61
Q96id9	homo sapien	277	4	Q96ID9
Q9v3f8	drosophila	280	5	Q9V3F8
Q96643	drosophila	280	5	Q96643
Q22841	arabidopsis	283	10	Q22841
Q91632	xenopus lae	292	13	Q91632
Q91633	xenopus lae	292	13	Q91633
Q8d3m7	vibrio vuln	294	16	Q8D3M7
Q8d3q3	vibrio vuln	299	16	Q8D3Q3
Q9a6s1	caulobacter	309	16	Q9A6S1
Q8nh68	homo sapien	312	4	Q8NH68
Q97d40	clostridium	326	16	Q97D40
Q91914	xenopus lae	327	13	Q91914
Q97d45	clostridium	350	16	Q97D45
Q9bye7	homo sapien	352	4	Q9BYE7
Q99na9	mus musculus	353	11	Q99NA9
Q90753	feline herp	371	12	Q90753
Q9aqp4	pseudomonas	386	2	Q9AQP4
Q8gi39	pseudomonas	386	2	Q8GI39
Q8rth6	vibrio chol	402	2	Q8RTH6

C 75 8 0.9 409 2 052274
 C 76 8 0.9 421 5 09XUV7
 C 77 8 0.9 424 2 093RK2
 C 78 8 0.9 428 2 069759
 C 79 8 0.9 437 4 08N7U5
 C 80 8 0.9 449 16 08EFM5
 C 81 8 0.9 475 16 092YEL
 C 82 8 0.9 485 5 09V9N1
 C 83 8 0.9 485 16 09CLJ7
 C 84 8 0.9 519 2 08RLN3
 C 85 8 0.9 521 2 093MK3
 C 86 8 0.9 524 2 093T64
 C 87 8 0.9 524 2 0938U2
 C 88 8 0.9 534 16 09EWX5
 C 89 8 0.9 544 16 09PMW3
 C 90 8 0.9 571 16 092BB4
 C 91 8 0.9 571 16 09Y6T8
 C 92 8 0.9 585 10 042880
 C 93 8 0.9 648 8 09TKV7
 C 94 8 0.9 664 16 08XL36
 C 95 8 0.9 671 6 09SLG6
 C 96 8 0.9 737 5 09TYX1
 C 97 8 0.9 758 17 09PW83
 C 98 8 0.9 761 4 08WUA9
 C 99 8 0.9 770 13 093394
 C 100 8 0.9 810 16 08ET53

ALIGNMENTS

RESULT 1
 Q8JIW9 PRELIMINARY; PRT; 774 AA.
 ID Q8JIW9; PRELIMINARY; PRT; 774 AA.
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Semaphorin 3A.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue-Head;
 RA Tannahill D., Nielsen J., Regan A.G.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AV030051; AAK38166.1; -;
 DR InterPro; IPR003599; Ig-like.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003006; Ig MHC.
 DR InterPro; IPR001005; MYB DNA binding.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS50835; IG LIKE; 1.
 DR PROSITE; PS00037; MYB 1; 1.
 SQ SEQUENCE 774 AA; 89319 MW; 3127FD2E0CCD5FA4 CRC64;

Alignment Scores:
 Pred. No.: 1.83e-37 Length: 774
 Score: 43.00 Matches: 43
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 4.80% Indels: 0
 DB: 13 Gaps: 0

US-09-774-490-1 (1-2709) x Q8JIW9 (1-774)
 QY 1004 TGCAGATGCTTGGAGGGCAGAGAGTCTGTGATTAATGACACATTCCTCAA 1063
 DB 270 CyslysAsnAspPheGlyGlyHisArgSerLeuValAsnLysTrpThrPheLeuLys 289
 QY 1064 GCTCGTCTGATTGCTCAGTCCAGGTCCTCAATGGCATTGACATCTTTGATGAAC 1123
 DB 290 AlaArgLeuLeuCysSerValProGlyProAsnGlyLeuAspThrHisPheAspGluLeu 309
 QY 1124 CAGGATGTA 1132
 DB 310 GlnAspVal 312

RESULT 2

Q9NS98 PRELIMINARY; PRT; 782 AA.
 ID Q9NS98; Q9H7Q3;
 AC Q9NS98; Q9H7Q3;
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Semaphorin sem2 (FLJ00014 protein).
 GN SEM2 OR FLJ00014.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Seki N., Hattori A., Hayashi A., Kozuma S., Muramatsu M., Miyajima N.,
 RA Saito T.;
 RT "Human semaphorin";
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 58-782 FROM N.A.
 RC Tissue-Spleen;
 RA Ohara O., Nagase T., Kikuno R., Okumura K.;
 RT "The nucleotide sequence of a long cDNA clone isolated from human spleen";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB029496; BAA98132.1; -;
 DR EMBL; AK024425; BAB15715.1; -;
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig-like.
 DR InterPro; IPR003006; Ig MHC.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF00047; Ig; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS50835; IG LIKE; 1.
 SQ SEQUENCE 782 AA; 86700 MW; 85CB424874DF6663 CRC64;

Alignment Scores:
 Pred. No.: 8.64e-09 Length: 782
 Score: 17.00 Matches: 17
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.90% Indels: 0
 DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q9NS98 (1-782)

QY 1766 GCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGTGGGTGGTGT 1816
 DB 528 AlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 544
 RESULT 3
 Q9HBR1 PRELIMINARY; PRT; 457 AA.
 ID Q9HBR1
 AC Q9HBR1;

RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB083186; BAB88870.1; --
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003006; Ig.MHC.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR001627; Sema.
DR Pfam; PF00047; Ig; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; Ig; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR PROSITE; PS0835; IG LIKE; 1.
SQ SEQUENCE 754 AA; 83691 MW; E61FD08C04E9A68E CRC64;
Alignment Scores:
Pred. No.: 1.4e-06 Length: 754
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
Gaps: 0
US-09-774-490-1 (1-2709) x Q8TDV7 (1-754)
Qy 1028 AGAAGCTCTGGTGAATAAATGGACAACATTCCTCAAGCTCGTCTG 1072
Db 277 ARGSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeu 291
RESULT 7
Q8BKQ6
ID Q8BKQ6 PRELIMINARY; PRT; 403 AA.
AC Q8BKQ6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Semaphorin 3D precursor homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Dorsal root ganglion;
EX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK051165; BAC34542.1; -- 1D5B284549825884 CRC64;
SQ SEQUENCE 403 AA; 45575 MW; 1D5B284549825884 CRC64;
Alignment Scores:
Pred. No.: 1.93e-05 Length: 403
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
Gaps: 0
US-09-774-490-1 (1-2709) x Q8BKQ6 (1-403)
Qy 1040 AATAAATGGACAACATTCCTCAAGCTCGTCTGATTGCTCA 1081
Db 298 AsnLysTrpThrPheLeuLysAlaArgLeuLysCysSer 311
RESULT 8
Q8BMF6
ID Q8BMF6 PRELIMINARY; PRT; 777 AA.
AC Q8BMF6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Semaphorin 3D precursor homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
EX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK031704; BAC27522.1; -- B7204D82288B89CE CRC64;
SQ SEQUENCE 777 AA; 89562 MW; B7204D82288B89CE CRC64;
Alignment Scores:
Pred. No.: 1.76e-05 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
Gaps: 0
US-09-774-490-1 (1-2709) x Q8BMF6 (1-777)
Qy 1040 AATAAATGGACAACATTCCTCAAGCTCGTCTGATTGCTCA 1081
Db 298 AsnLysTrpThrPheLeuLysAlaArgLeuLysCysSer 311
RESULT 9
Q8BH34
ID Q8BH34 PRELIMINARY; PRT; 777 AA.
AC Q8BH34;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Semaphorin 3D precursor homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Kidney, and Skin;
EX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK028900; BAC26185.1; --
DR EMBL; AK052671; BAC35092.1; --
SQ SEQUENCE 777 AA; 89548 MW; 5450D8D45D1BDABF CRC64;
Alignment Scores:
Pred. No.: 1.76e-05 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
Gaps: 0
US-09-774-490-1 (1-2709) x Q8BH34 (1-777)
Qy 1040 AATAAATGGACAACATTCCTCAAGCTCGTCTGATTGCTCA 1081
Db 298 AsnLysTrpThrPheLeuLysAlaArgLeuLysCysSer 311
RESULT 10
Q9VHX3

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Semaphorin M-Semak.
 GN SEMA3E.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57/Black 6;
 RA Miyazaki N., Furiyama T., Inagaki S.;
 RT "A novel semaphorin, M-Semak which inhibits neural outgrowth from
 RT sensory neurons."
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF034744; RAD01996.1; -
 DR MGD; MGI:1340034; Semak3e.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003006; IG.MHC.
 DR InterPro; IPR003659; plexin-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF00047; ig; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS50835; IG_LIKE; 1.
 SQ SEQUENCE 775 AA; 89543 MW; 221E766FF404098D4 CRC64;

Alignment Scores:
 Pred. No.: 0.000223 Length: 775
 Score: 13.00 Matches: 13
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.45% Indels: 0
 DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x Q90X23 (1-775)

QY 1778 TGTTCCTCGCCGAGACCTTACTGTCTGGATGGT 1816
 Db 529 CysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 541

RESULT 14
 Q8NAS4 PRELIMINARY; PRT; 132 AA.
 ID Q8NAS4;
 AC Q8NAS4;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein FLJ34863.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
 RA Furiyama T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamiyama K.,
 RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
 RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
 RA Matsumoto K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
 RA Kanchori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
 RA Suzuki Y., Sugano S., Nagahara K., Masuno Y., Nagai K., Isogai T.;
 RT "NEDO human cDNA sequencing project."
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK092182; BAC03823.1; -
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 132 AA; 14227 MW; BF1D942742DE0FA1 CRC64;

Alignment Scores:
 Pred. No.: 0.00163 Length: 132
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.34% Indels: 0
 DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q8NAS4 (1-132)

QY 1547 TTTATCGAACAGATGTTGGACCGTCTTAAAGTA 1582
 Db 42 PheileGlyThrAspValGlyThrValLeuLysVal 53

RESULT 15
 Q9J129 PRELIMINARY; PRT; 296 AA.
 ID Q9J129;
 AC Q9J129;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Semaphorin (Fragment).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley;
 RA Brenz Verca M.S., Widmer D.A.J., Wagner G.C., Dreyer J.L.;
 RT "New rat semaphorin with closest homology to avian Sema3D (C-
 RT collapsein-2)."
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF685594; AAF76329.1; -
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00630; Sema; 1.
 DR NON_TER 1
 FT NON_TER 296
 SQ SEQUENCE 296 AA; 33558 MW; C2226212885D044C CRC64;

Alignment Scores:
 Pred. No.: 0.00324 Length: 296
 Score: 12.00 Matches: 12
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.34% Indels: 0
 DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x Q9J129 (1-296)

QY 1319 CCCTATCCAGCCAGCACTGTCCAGCAACACA 1354
 Db 136 ProTyrProArgProGlyThrCysProSerLysThr 147

RESULT 16
 Q8AXA8 PRELIMINARY; PRT; 160 AA.
 ID Q8AXA8;
 AC Q8AXA8;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Secreted chemorepellent semaphorin 3F (Fragment).
 GN SEMA3F.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chilton J.K., Guthrie S.;
 RT "Cranial expression of semaphorins and neuropilins in the embryonic

RT chick.";
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF495525; AAN77007.1; --
 FT NON TER 160
 SQ SEQUENCE 160 AA; 17865 MW; 8033FOC191EAF2DE CRC64;

Alignment Scores:
 Pred. No.: 0.567 Length: 160
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.12% Indels: 0
 DB: 13 Gaps: 0

US-09-774-490-1 (1-2709) x Q9AXA8 (1-160)

QY 1106 ACTATTGATGAGTCTGATCTATTC 1135
 Db 108 ThrHisPheAspGluLeuGlnAspValPhe 117

RESULT 17
 Q64906 PRELIMINARY; PRT; 653 AA.
 AC Q64906;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to GENBANK ACCESSION number L26081.
 OS Alcelaphine herpesvirus 1 (wildbeest herpesvirus).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 OX NCBI_TaxID=35252;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C500;
 RX MEDLINE=97201573; PubMed=9049359;
 RA Enser A., Fleckenstein B.;
 RT "Alcelaphine herpesvirus type 1 has a semaphorin-like gene.";
 RL J. Gen. Virol. 76:1063-1067(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C500;
 RX MEDLINE=97404659; PubMed=9261371;
 RA Enser A., Pflanz R., Fleckenstein B.;
 RT "Primary structure of the alcelaphine herpesvirus 1 genome.";
 RL J. Virol. 71:6517-6525(1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C500;
 RA Enser A., Pflanz R., Fleckenstein B.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U18243; AAC54475.1; --
 DR EMBL; AF005370; AAC58054.1; --
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS0835; IG_LIKE; 1.
 SQ SEQUENCE 653 AA; 73645 MW; 5C2816BD567E706E CRC64;

Alignment Scores:
 Pred. No.: 0.466 Length: 653
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.12% Indels: 0
 DB: 12 Gaps: 0

US-09-774-490-1 (1-2709) x Q64906 (1-653)

QY 1043 AAATGACACATTTCTCAAGCTGCTG 1072
 Db 302 LysTrpThrPheLeuLysAlaArgLeu 311

RESULT 18
 Q9C0C4 PRELIMINARY; PRT; 963 AA.
 AC Q9C0C4;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein KIAA1739 (Fragment).
 GN KIAA1739.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21082932; PubMed=11214970;
 RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XIX.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro.";
 RL DNA Res. 7:347-355(2000).
 DR EMBL; AB051526; BAB21830.1; --
 DR Genew; HGNC:10731; SEMA4C.
 DR InterPro; IPR003599; IG.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR002165; Plexin_repeat.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01437; PSI; 1.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR SMART; SM00630; Sema; 1.
 DR PROSITE; PS0835; IG_LIKE; 1.
 KW Hypothetical protein.
 FT NON TER 1
 SQ SEQUENCE 963 AA; 106735 MW; 3FA37DBBA3483ECE CRC64;

Alignment Scores:
 Pred. No.: 0.441 Length: 963
 Score: 10.00 Matches: 10
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.12% Indels: 0
 DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q9C0C4 (1-963)

QY 917 GATGACAAAGTATCTTTTCTCCGTGAA 946
 Db 362 AspaAspLysValTyrPhePheArgGlu 371

RESULT 19
 Q96QN2 PRELIMINARY; PRT; 139 AA.
 AC Q96QN2;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Complement component 1, s subcomponent.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 CUB DOMAIN.
 DR EMBL; BC007010; AAH07010.1; --
 DR InterPro; IPR000859; CUB_domain.

DR Pfam; PF00431; CUB; 1.
 DR SMART; SM00042; CUB; 1.
 DR PROSITE; PS01180; CUB; 1.
 SQ SEQUENCE 139 AA; 15969 MW; 590D59BC718112D5 CRC64;

Alignment Scores:
 Pred. No.: 7.33 Length: 139
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.03% Indels: 0
 DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q96QN2 (1-139)

QY 151 TTGTGGCGCCAGAGAGTTCAACAT 125
 Db 82 LeuCysGlyGlnArgSerSerAsnAsn 90

RESULT 20

Q49704 PRELIMINARY; PRT; 223 AA.
 AC Q49704;
 DT 01-NOV-1996 (TReMBLrel. 01, Created)
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
 DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE B1496_F2_61.
 OS Mycobacterium leprae.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Smith D.R.;
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Robison K.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U00013; AA017144.1; -.
 DR InterPro; IPR003780; COX15_CtaA.
 DR Pfam; PF02628; COX15-CtaA; 1.
 SQ SEQUENCE 223 AA; 23338 MW; 42D301144DF0B355 CRC64;

Alignment Scores:
 Pred. No.: 6.86 Length: 223
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.03% Indels: 0
 DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x Q49704 (1-223)

QY 516 GCCACTTGCAATTCATCTTCTGTG 490
 Db 137 AlaHisLeuHisSerSerLeuVal 145

RESULT 21

O54948 PRELIMINARY; PRT; 284 AA.
 AC O54948;
 DT 01-JUN-1998 (TReMBLrel. 06, Created)
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)
 DE 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
 DE Senaphorin IV homolog (Fragment).
 GN SEMA3F.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Hu L.-J., Drabkin H.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF036162; AAB88861.1; -.
 DR MGI; MGI:1096347; Sema3f.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 DR SMART; SM00630; Sema; 1.
 FT NON TER 284 284
 SQ SEQUENCE 284 AA; 32151 MW; BFD8D7E66BD08027 CRC64;

Alignment Scores:
 Pred. No.: 6.63 Length: 284
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.01% Indels: 0
 DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x O54948 (1-284)

QY 1106 ACTCATTTTGATGAACGACGATGTA 1132
 Db 235 ThrHisPheAspGluLeuGlnAspVal 243

RESULT 22

O33142 PRELIMINARY; PRT; 311 AA.
 ID O33142
 AC O33142;
 DT 01-JAN-1998 (TReMBLrel. 05, Created)
 DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Hypothetical 32.6 kDa protein (Putative membrane protein).
 GN ML0587 OR MLCL536.34.
 OS Mycobacterium leprae.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Seeger K., Harris D.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TN;
 RC MEDLINE=21128732; PubMed=11234002;
 RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
 RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Barrell B.G.;
 RT "Massive gene decay in the leprosy bacillus.";
 RL Nature 409:1007-1011(2001).
 DR EMBL; Z99125; CAB16178.1; -.
 DR EMBL; AL583919; CAC30095.1; -.
 DR Leproma; ML0587; -.
 DR InterPro; IPR003780; COX15_CtaA.
 DR Pfam; PF02628; COX15-CtaA; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 311 AA; 32590 MW; 8F5B91AAA6378E08 CRC64;

Alignment Scores:
 Pred. No.: 6.55 Length: 311
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.03% Indels: 0
 DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x O33142 (1-311)

QY 516 GCCCAGCTTGCATTCATCTCTCTCTGCTG 490
 |||||
 Db 225 AlahieLeuHieserSerLeuLeuVal 233

RESULT 23
 Q8BW98
 ID Q8BW98 PRELIMINARY; PRT; 325 AA.
 AC Q8BW98;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Similar to eukaryotic translation initiation factor 3, subunit 8 (110kD).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Muscle;
 RA Strausberg R.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC000533; AA00533.1; -.
 DR InterPro; IPR000717; FCI.
 DR Pfam; PF01399; PCI; 1.
 DR SMART; SM00088; PINT; 1.
 KW Initiation factor.
 SQ SEQUENCE 325 AA; 37680 MW; 701BESA01DD7E10B CRC64;

Alignment Scores:
 Pred. No.: 6.51 Length: 325
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.01% Indels: 0
 DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q8BW98 (1-325)

QY 2093 CAAGCCTTCGTCTAGCTACAA 2119
 |||||
 Db 60 GlnGlyLeuLeuArgSerLeuGln 68

RESULT 24
 Q8NEN9
 ID Q8NEN9 PRELIMINARY; PRT; 333 AA.
 AC Q8NEN9;
 DT 01-OCT-2002 (TREMBLrel. 22, Created)
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Hypothetical protein NT2RP2004036.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Ota T., Nishikawa T., Suzuki Y., Kawai-Hio Y., Hayashi K., Ishii S.,
 RA Saito K., Yamamoto J., Wakamatsu A., Nagai T., Nakamura Y.,
 RA Nagahari K., Sugano S., Isogai T.;
 RT "HRI human cDNA sequencing project."
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK075388; BAC11588.1; -.
 DR InterPro; IPR003599; IG-like.
 DR InterPro; IPR007110; IG-like.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR002165; Plexin-like.
 DR Pfam; PF01437; PSI; 1.
 DR SMART; SM00409; IG; 1.
 DR SMART; SM00423; PSI; 1.
 DR PROSITE; PS50835; IG LIKE; 1.
 KW Hypothetical protein.

SQ SEQUENCE 333 AA; 35797 MW; F74AF2628FA274F3 CRC64;

Alignment Scores:
 Pred. No.: 6.48 Length: 333
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.01% Indels: 0
 DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q8NEN9 (1-333)

QY 1784 CTCGCCGAGACCTTACTGTGCTTG 1810
 |||||
 Db 11 LeuAlaArgAspProTyrCysAlaTir 19

RESULT 25
 Q8F7R6
 ID Q8F7R6 PRELIMINARY; PRT; 350 AA.
 AC Q8F7R6;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Dna protein.
 GN DSHA OR LA0878.
 OS Leptospira interrogans.
 OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
 OX NCBI_TaxID=173;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
 RA Ren S.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AE011272; AAN48077.1; -.
 KW Complete proteome.
 SQ SEQUENCE 350 AA; 39815 MW; 78044D56923114CD CRC64;

Alignment Scores:
 Pred. No.: 6.44 Length: 350
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.01% Indels: 0
 DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x Q8F7R6 (1-350)

QY 798 CTATCTTCGAACTCTTGGCACCAC 824
 |||||
 Db 188 LeuSerSerGluLeuLeuGlyThrThr 196

RESULT 26
 Q8WVP9
 ID Q8WVP9 PRELIMINARY; PRT; 458 AA.
 AC Q8WVP9;
 DT 01-MAR-2002 (TREMBLrel. 20, Created)
 DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
 DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
 DE Hypothetical protein (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RA Strausberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC017658; AA017658.1; -.
 DR InterPro; IPR003659; Plexin-like.
 DR InterPro; IPR002165; Plexin-repeat.
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01437; PSI; 1.

```
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00423; PSI; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 458 AA; 50263 MW; B8A52233DCE848E7 CRC64;

Alignment Scores:
Pred. No.: 6.2 Length: 458
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q8WVP9 (1-458)

QY 1784 CTCGCCGAGACCTTACTGTCTGG 1810
Db 158 LeuAlaArgAspProTyrCysAlaTrp 166

RESULT 27
Q96FK5 PRELIMINARY; PRT; 475 AA.
AC Q96FK5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC010701; AAH10701.1; -.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002165; Plexin_repeat.
DR InterPro; IPR001627; Sema.
DR Pfam; PF01437; PSI; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00423; PSI; 1.
KW Hypothetical protein.
SQ SEQUENCE 475 AA; 52298 MW; EE90909701AE2264 CRC64;

Alignment Scores:
Pred. No.: 6.17 Length: 475
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q96FK5 (1-475)

QY 1784 CTCGCCGAGACCTTACTGTCTGG 1810
Db 175 LeuAlaArgAspProTyrCysAlaTrp 183

RESULT 28
Q9NX92 PRELIMINARY; PRT; 510 AA.
AC Q9NX92;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein FLJ20369.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
```

```
[1]
RN SEQUENCE FROM N.A.
RP Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
RA Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Obayashi M.,
RA Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK000376; BAA91124.1; -.
DR InterPro; IPR003599; Ig-like.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002165; Plexin_repeat.
DR InterPro; IPR001627; Sema.
DR Pfam; PF01437; PSI; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00409; IG; 1.
DR SMART; SM00423; PSI; 1.
DR PROSITE; PS00835; IG_LIKE; 1.
KW Hypothetical protein.
SQ SEQUENCE 510 AA; 56147 MW; 53125912E9AD0F04 CRC64;

Alignment Scores:
Pred. No.: 6.11 Length: 510
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q9NX92 (1-510)

QY 1784 CTCGCCGAGACCTTACTGTCTGG 1810
Db 188 LeuAlaArgAspProTyrCysAlaTrp 196

RESULT 29
Q8BIC3 PRELIMINARY; PRT; 550 AA.
AC Q8BIC3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Semaphorin 4B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK028896; BAC26181.1; -.
SQ SEQUENCE 550 AA; 61309 MW; 8C28E29DC4CAC958 CRC64;

Alignment Scores:
Pred. No.: 6.04 Length: 550
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x Q8BIC3 (1-550)

QY 1784 CTCGCCGAGACCTTACTGTCTGG 1810
Db 524 LeuAlaArgAspProTyrCysAlaTrp 532
```

```
RESULT 30
QBUC1
ID Q8BJC1 PRELIMINARY; PRT; 799 AA.
AC Q8BJC1;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Semaphorin 4D precursor (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOB; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK088653; BAC40480.1; -.
FT NON_TER 1
SQ SEQUENCE 799 AA; 88814 MW; 415D3B687150A59A CRC64;

Alignment Scores:
Pred. No.: 5.74 Length: 799
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x Q8BJC1 (1-799)

QY 914 GAGATGACAAAGTACTTTTCCTC 940
Db 165 GluAspAspLysValTyrPhePhe 173

RESULT 31
QBIR6
ID Q8BIR6 PRELIMINARY; PRT; 823 AA.
AC Q8BIR6;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Semaphorin 4B
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Ovary, and Uterus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium.
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK030331; BAC26906.1; -.
SQ SEQUENCE 823 AA; 91485 MW; E3B8C5E5CFA1CD45 CRC64;

Alignment Scores:
Pred. No.: 5.71 Length: 823
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x Q8BIR6 (1-823)
```

```
QY 1784 CTCGCCGAGACCTTACTGTGTGG 1810
Db 524 LeuAlaArgAspProTyrCysAlaTyr 532

RESULT 32
Q9C0B8
ID Q9C0B8 PRELIMINARY; PRT; 893 AA.
AC Q9C0B8;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical protein KIAA1745 (Fragment).
GN KIAA1745.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21082932; PubMed=11214970;
RA Nagase T., Kikuno R., Hattori A., Kondo Y., Okumura K., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIX.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 7:347-355 (2000).
DR EMBL; AB051532; BAB21836.1; -.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002165; Plexin_repeat.
DR InterPro; IPR001627; Sema.
DR Pfam; PF01437; PSI; 1.
DR Pfam; PF01403; Sema; 1.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 893 AA; 98361 MW; D6C6C48DEE524F14 CRC64;

Alignment Scores:
Pred. No.: 5.65 Length: 893
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q9C0B8 (1-893)

QY 1784 CTCGCCGAGACCTTACTGTGTGG 1810
Db 593 LeuAlaArgAspProTyrCysAlaTyr 601

RESULT 33
Q8R3M7
ID Q8R3M7 PRELIMINARY; PRT; 911 AA.
AC Q8R3M7;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Similar to eukaryotic translation initiation factor 3, subunit 8
DE (110kD).
GN Mus musculus (Mouse).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Strausberg R.;
RA Submitted (MAR-2002) to the EMBL/GenBank/DBSJ databases.
RL EMBL; BC025032; AAH25032.1; -.
DR MGD; MGI:1913935; 3230401013Rik.
DR InterPro; IPR001395; Aldo/ket_red.
DR InterPro; IPR000717; PCI.
```

DR Pfam; PF01399; PCI; 1.
 DR SMART; SM00088; PINT; 1.
 DR PROSITE; PS00063; ALDOXETO_REDUCTASE_3; 1.
 KW Initiation factor.
 SQ SEQUENCE 911 AA; 105545 MW; 5A2D29E961073084 CRC64;

Alignment Scores:
 Pred. No.: 5.63 Length: 911
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.01% Indels: 0
 DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x Q8R3M7 (1-911)

QY 2093 CAAGGCTTCGCTAGTAGTCTACAA 2119
 Db 646 GlnGlyLeuLeuArgSerLeuGln 654

RESULT 34
 Q8R1B4 PRELIMINARY; PRT; 911 AA.
 ID Q8R1B4
 AC Q8R1B4;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Similar to eukaryotic translation initiation factor 3, subunit 8 (110KD).
 DE (110KD).
 GN 3230401013RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Eye;
 RA Strausberg R.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBSJ databases.
 DR EMBL; BC024855; AAH24855.1;
 DR MGI; MGI:1913935; 3230401013RIK.
 DR InterPro; IPR001395; Aldo/ket_red.
 DR InterPro; IPR000717; PCI.
 DR SMART; SM00088; PINT; 1.
 DR PROSITE; PS00063; ALDOXETO_REDUCTASE_3; 1.
 KW Initiation factor.
 SQ SEQUENCE 911 AA; 105531 MW; A88282C5B4C55A8B CRC64;

Alignment Scores:
 Pred. No.: 5.63 Length: 911
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.01% Indels: 0
 DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x Q8R1B4 (1-911)

QY 2093 CAAGGCTTCGCTAGTAGTCTACAA 2119
 Db 646 GlnGlyLeuLeuArgSerLeuGln 654

RESULT 35
 Q8R1B4 PRELIMINARY; PRT; 4859 AA.
 ID Q8R1B4
 AC Q8R1B4;
 DT 01-FEB-1997 (TrEMBLrel. 02, Created)
 DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Ryanodine receptor type 3.
 GN RYR3.
 OS Muscivora vison (American mink).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Mustelidae; Mustelinae;
 OC Mustela.
 OX NCBI_TaxID=9667;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Marziani G., Rossi D., Giannini G., Charlesworth A., Sorrentino V.;
 RT "cDNA cloning reveals a tissue specific expression of alternatively
 RT spliced transcripts of the ryanodine receptor type 3 (RYR3) calcium
 RT release channel";
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBSJ databases.
 DR EMBL; Y07749; CAA69029.1;
 DR InterPro; IPR000194; ATPase_a/bcentre.
 DR InterPro; IPR000699; Ca-rel_channel.
 DR InterPro; IPR001682; Ca/Na_Pore.
 DR InterPro; IPR005797; Cytb_56_N.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR005821; Ion_trans.
 DR InterPro; IPR003608; MIR.
 DR InterPro; IPR001215; Ryanodn_receptor.
 DR InterPro; IPR003032; RyR.
 DR InterPro; IPR003877; SPRY_receptor.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF00520; Ion_trans; 1.
 DR Pfam; PF02815; MIR; 4.
 DR Pfam; PF01365; RYDR_ITPR; 2.
 DR Pfam; PF02026; RYR; 4.
 DR Pfam; PF00622; SPRY; 3.
 DR PRINTS; PR00795; RYANODINER.
 DR SMART; SM00472; MIR; 4.
 DR SMART; SM00449; SPRY; 3.
 DR PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 KW Ionic channel; Receptor; Transmembrane.
 SQ SEQUENCE 4859 AA; 550457 MW; B3B5EC108C3A653D CRC64;

Alignment Scores:
 Pred. No.: 4.46 Length: 4859
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservatives: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.03% Indels: 0
 DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x Q95201 (1-4859)

QY 67 CGTGGTCGACACCCCTCGACGTGGCA 41
 Db 2132 ArgGlySerThrProLeuAspValAla 2140

RESULT 36
 Q91319 PRELIMINARY; PRT; 4868 AA.
 ID Q91319
 AC Q91319;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE Ryanodine receptor beta isoform.
 GN FROG-BRR.
 OS Rana catesbeiana (Bull frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Neobatrachia; Ranioidea; Rana.
 OX NCBI_TaxID=8400;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=94274714; PubMed=8006029;
 RA Oyama H., Murayama T., Takagi T., Iino M., Iwabe N., Miyata T.,
 RA Oyama Y., Endo M.;
 RT "Primary structure and distribution of ryanodine-binding protein
 RT isoforms of the bullfrog skeletal muscle.";
 RL J. Biol. Chem. 269:17206-17214(1994).
 DR EMBL; D21071; BAA04647.1;
 DR InterPro; IPR000699; Ca-rel_channel.

DR InterPro; IPR001682; Ca/Na_pore.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR005821; Ion_trans.
 DR InterPro; IPR003608; MIR.
 DR InterPro; IPR001215; Ryanodn_receptor.
 DR InterPro; IPR003032; RYR.
 DR InterPro; IPR003877; SPRY_receptor.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF00520; ion_trans; 1.
 DR Pfam; PF02815; MIR; 4.
 DR Pfam; PF01365; RYDR_ITPR; 2.
 DR Pfam; PF02026; RYR; 4.
 DR Pfam; PF00622; SPRY; 3.
 DR PRINTS; PR00795; RYANODINER.
 DR SMART; SM00472; MIR; 4.
 DR SMART; SM00449; SPRY; 3.
 KW Ionic channel; Receptor; Transmembrane.
 SQ SEQUENCE 4868 AA; 553045 MW; F5187BD678C369D1 CRC64;

Alignment Scores:
 Pred. No.: 4.46 Length: 4868
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.03% Indels: 0
 DB: 13 Gaps: 0
 US-09-774-490-1 (1-2709) x Q91319 (1-4868)
 QY 67 CGTGGTGCACACCCCTCGACGTGCA 41
 Db 2136 ArgGlySerThrProLeuAspValala 2144

RESULT 37
 Q90985 PRELIMINARY; PRT; 4869 AA.
 AC Q90985;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Ryanodine receptor type 3.
 GN RYR3.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96207583; PubMed=8670108;
 RA Ottini L., Marzali G., Conti A., Charlesworth A., Sorrentino V.;
 RT "alpha and beta isoforms of ryanodine receptor from chicken skeletal
 muscle are the homologues of mammalian RYR1 and RYR3.";
 RL Biochem. J. 315:207-216(1996).
 DR EMBL; X95267; CAA64563.1; -;
 DR InterPro; IPR000699; Ca-rel_channel.
 DR InterPro; IPR001682; Ca/Na_pore.
 DR InterPro; IPR005797; Cytb_b6_N.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR005821; Ion_trans.
 DR InterPro; IPR001215; Ryanodn_receptor.
 DR InterPro; IPR003032; RYR.
 DR InterPro; IPR003877; SPRY_receptor.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF00520; ion_trans; 1.
 DR Pfam; PF02815; MIR; 4.
 DR Pfam; PF01365; RYDR_ITPR; 2.
 DR Pfam; PF02026; RYR; 4.
 DR Pfam; PF00622; SPRY; 3.
 DR PRINTS; PR00795; RYANODINER.
 DR SMART; SM00472; MIR; 4.
 DR SMART; SM00449; SPRY; 3.
 DR PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 KW Ionic channel; Receptor; Transmembrane.
 SQ SEQUENCE 4872 AA; 551925 MW; D1DC6B85CD6B93E6 CRC64;

Alignment Scores:
 Pred. No.: 4.45 Length: 4872
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.03% Indels: 0
 DB: 6 Gaps: 0
 US-09-774-490-1 (1-2709) x Q9TS33 (1-4872)

DR SMART; SM00449; SPRY; 3.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 KW Ionic channel; Receptor; Transmembrane.
 SQ SEQUENCE 4869 AA; 552767 MW; E04D6260F20FFA21 CRC64;
 Alignment Scores:
 Pred. No.: 4.46 Length: 4869
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.03% Indels: 0
 DB: 13 Gaps: 0
 US-09-774-490-1 (1-2709) x Q90985 (1-4869)
 QY 67 CGTGGTGCACACCCCTCGACGTGCA 41
 Db 2135 ArgGlySerThrProLeuAspValala 2143

RESULT 38
 Q9TS33 PRELIMINARY; PRT; 4872 AA.
 AC Q9TS33;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE RYANODINE receptor/calcium release channel.
 OS Oryctolagus sp.; and
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=60556; 9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93050200; PubMed=1330694;
 RA Hakamata Y., Nakai J., Takeshima H., Imoto K.;
 RT "Primary structure and distribution of a novel ryanodine
 receptor/calcium release channel from rabbit brain.";
 RL FEBS Lett. 312:229-235(1992).
 DR EMBL; X68650; CAC16153.1; -;
 DR InterPro; IPR000194; ATPase_a/bcentre.
 DR InterPro; IPR000699; Ca-rel_channel.
 DR InterPro; IPR001682; Ca/Na_pore.
 DR InterPro; IPR005797; Cytb_b6_N.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR005821; Ion_trans.
 DR InterPro; IPR003608; MIR.
 DR InterPro; IPR001215; Ryanodn_receptor.
 DR InterPro; IPR003032; RYR.
 DR InterPro; IPR003877; SPRY_receptor.
 DR Pfam; PF00036; ehand; 1.
 DR Pfam; PF00520; ion_trans; 1.
 DR Pfam; PF02815; MIR; 4.
 DR Pfam; PF01365; RYDR_ITPR; 2.
 DR Pfam; PF02026; RYR; 4.
 DR Pfam; PF00622; SPRY; 3.
 DR PRINTS; PR00795; RYANODINER.
 DR SMART; SM00472; MIR; 4.
 DR SMART; SM00449; SPRY; 3.
 DR PROSITE; PS00152; ATPASE_ALPHA_BETA; 1.
 DR PROSITE; PS00192; CYTOCHROME_B_HEME; 1.
 KW Ionic channel; Receptor; Transmembrane.
 SQ SEQUENCE 4872 AA; 551925 MW; D1DC6B85CD6B93E6 CRC64;

QY 67 CGTGGGTCGACACCCCTCGACGTGGCA 41
DB 2136 ArgGlySerThrProLeuAspValAla 2144

RESULT 39
Q9ERN6
ID Q9ERN6 PRELIMINARY; PRT; 4967 AA.
AC Q9ERN6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Cardiac Ca2+ release channel.
GN 9330127120RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99403027; PubMed=10473538;
RA Zhao M., Li P., Li X., Zhang L., Winkfein R.J., Chen S.R.;
RT "Molecular identification of the ryanodine receptor pore-forming
segment.";
RL J. Biol. Chem. 274:25971-25974(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Chen S.W.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF295105; AAG34081.1; -;
DR MGD; MG1:1924803; 9330127120RIK.
DR InterPro; IPR000699; Ca-rel_channel.
DR InterPro; IPR001682; Ca/Na_pore.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR005821; Ion trans.
DR InterPro; IPR003608; MIR.
DR InterPro; IPR001215; Ryanodn_receptor.
DR InterPro; IPR003032; RYR.
DR InterPro; IPR003877; SPRY_receptor.
DR Pfam; PF00036; ehand; 2.
DR Pfam; PF00520; ion_trans; 1.
DR Pfam; PF02815; MIR; 4.
DR Pfam; PF01365; RYDR_ITPR; 2.
DR Pfam; PF02026; RYR; 4.
DR PRINTS; PR00795; RYANODINER.
DR SMART; SM00472; MIR; 4.
DR SMART; SM00449; SPRY; 3.
DR SMART; SM00449; Transmembrane.
KW Ionic channel; Transmembrane.
SQ SEQUENCE 4967 AA; 564887 MW; 81A27F4DCFF4E4FB CRC64;

Alignment Scores:
Pred. No.: 4.44 Length: 4967
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x Q9ERN6 (1-4967)

QY 67 CGTGGGTCGACACCCCTCGACGTGGCA 41
DB 2234 ArgGlySerThrProLeuAspValAla 2242

RESULT 40
Q29621
ID Q29621 PRELIMINARY; PRT; 4968 AA.
AC Q29621;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Cardiac RYANODINE receptor.

GN RYR-2.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart, and Brain;
RX MEDLINE=91032060; PubMed=2226801;
RA Nakai J., Imagawa T., Hakamat Y., Shigekawa M., Takeshima H., Numa S.;
RT "Primary structure and functional expression from cDNA of the cardiac
ryanodine receptor/calcium release channel.";
RL FEBS Lett. 271:169-177(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart, and Brain;
RX MEDLINE=93147068; PubMed=1337083;
RA Imagawa T., Nakai J., Takeshima H., Nakasaki Y., Shigekawa M.;
RT "Expression of Ca(2+)-induced Ca2+ release channel activity from
cardiac ryanodine receptor cDNA in Chinese hamster ovary cells.";
RL J. Biochem. 112:508-513(1992).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart, and Brain;
RA Nakai J.;
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; U50465; AAA93465.1; -;
DR EMBL; U50466; AAA93469.1; -;
DR InterPro; IPR000699; Ca-rel_channel.
DR InterPro; IPR001682; Ca/Na_pore.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR005821; Ion trans.
DR InterPro; IPR003608; MIR.
DR InterPro; IPR001215; Ryanodn_receptor.
DR InterPro; IPR003032; RYR.
DR InterPro; IPR003877; SPRY_receptor.
DR Pfam; PF00036; ehand; 2.
DR Pfam; PF00520; ion_trans; 1.
DR Pfam; PF02815; MIR; 4.
DR Pfam; PF01365; RYDR_ITPR; 2.
DR Pfam; PF02026; RYR; 4.
DR PRINTS; PR00795; RYANODINER.
DR SMART; SM00472; MIR; 4.
DR SMART; SM00449; SPRY; 3.
DR SMART; SM00449; Transmembrane.
KW Ionic channel; Transmembrane.
SQ SEQUENCE 4968 AA; 565081 MW; 04B16A35D06B852D CRC64;

Alignment Scores:
Pred. No.: 4.44 Length: 4968
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x Q29621 (1-4968)

QY 67 CGTGGGTCGACACCCCTCGACGTGGCA 41
DB 2235 ArgGlySerThrProLeuAspValAla 2243

RESULT 41
Q91313
ID Q91313 PRELIMINARY; PRT; 5037 AA.
AC Q91313;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Ryanodine receptor alpha isoform.
GN FROG-ARR.
OS Rana catesbeiana (Bull frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RX MEDLINE=94274714; PubMed=8006029;
RA Oyama H., Murayama T., Takagi T., Iino M., Iwabe N., Miyata T.,
Ogawa Y., Endo M.;
RT "Primary structure and distribution of ryanodine-binding protein
isoforms of the bullfrog skeletal muscle.";
RL J. Biol. Chem. 269:17206-17214(1994).
DR EMBL: D21070; BA04646.1; -;
DR InterPro: IPR000699; Ca-rel_channel.
DR InterPro: IPR001682; Ca/Na_pore.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR005821; Ion_trans.
DR InterPro: IPR003608; MIR.
DR InterPro: IPR001215; Ryanodn_receptor.
DR InterPro: IPR003032; Ryr.
DR InterPro: IPR003877; SPRY_receptor.
DR Pfam: PF00520; ion_trans; 1.
DR Pfam: PF02815; MIR; 4.
DR Pfam: PF01365; RYDR_ITPR; 2.
DR Pfam: PF02026; RYR; 4.
DR Pfam: PF00622; SPRY; 3.
DR PRINTS; PR00795; RYANODINER.
DR SMART; SM00472; MIR; 4.
DR SMART; SM00449; SPRY; 3.
KW Ionic channel; Receptor; Transmembrane.
SQ SEQUENCE 5037 AA; 571278 MW; FF263948A51C3ABF CRC64;

Alignment Scores:
Pred. No.: 4.43 Length: 5037
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 13 Gaps: 0

US-09-774-490-1 (1-2709) x Q91313 (1-5037)
QY 67 CGTGGTGCACACCCCTCGACGTGCA 41
|||||
DB 2253 ArgGlySerThrProLeuAspValala 2261

RESULT 42
O13054 PRELIMINARY; PRT; 5081 AA.
AC O13054;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Ryanodine receptor RYR1 isoform.
GN RYR1.
OS Makaira nigricans (Blue marlin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Scombroidei;
OC Xiphidae; Makaira.
OX NCBI_TaxID=13604;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skeletal muscle;
RA Franck J.P.C., Keen J.E., Londraville R.L., Morrisette J.,
Beamsley M., Block B.A.;
RT "Cloning and characterization of fiber type-specific ryanodine
receptor isoforms in skeletal muscles of fish (Makaira nigricans).";
RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U97329; AAB58117.1; -;
DR InterPro: IPR000699; Ca-rel_channel.
DR InterPro: IPR001682; Ca/Na_pore.
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR005821; Ion_trans.

DR InterPro: IPR003608; MIR.
DR InterPro: IPR001215; Ryanodn_receptor.
DR InterPro: IPR003032; RYR.
DR InterPro: IPR003877; SPRY_receptor.
DR Pfam: PF00520; ion_trans; 1.
DR Pfam: PF02815; MIR; 4.
DR Pfam: PF01365; RYDR_ITPR; 2.
DR Pfam: PF02026; RYR; 4.
DR Pfam: PF00622; SPRY; 3.
DR PRINTS; PR00795; RYANODINER.
DR SMART; SM00472; MIR; 4.
DR SMART; SM00449; SPRY; 3.
KW Ionic channel; Receptor; Transmembrane.
SQ SEQUENCE 5081 AA; 576220 MW; 5A3D7E253CFBC09B CRC64;

Alignment Scores:
Pred. No.: 4.43 Length: 5081
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 13 Gaps: 0

US-09-774-490-1 (1-2709) x O13054 (1-5081)
QY 67 CGTGGTGCACACCCCTCGACGTGCA 41
|||||
DB 2306 ArgGlySerThrProLeuAspValala 2314

RESULT 43
Q24498 PRELIMINARY; PRT; 5107 AA.
ID Q24498
AC Q24498; Q24500; Q24501; Q9V4Y7;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Ryanodine receptor.
GN RYA-R44F OR CG10844.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 2; 3; 4 AND 5), AND CHARACTERIZATION.
RX MEDLINE=94102409; PubMed=8276118;
RA Takeshima H., Nishi M., Iwabe N., Miyata T., Hosoya T., Masai I.,
Hotta Y.;
RT "Isolation and characterization of a gene for a ryanodine
receptor/calcium release channel in Drosophila melanogaster.";
RL FEBS Lett. 337:81-87(1994).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton R.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
Bailey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush P., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobaric C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Murny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reiner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kianos I., Simpson M., Skupek M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
CC -!- FUNCTION: COMMUNICATION BETWEEN TRANSVERSE-TUBULES AND
CC SARCOPLASMIC RETICULUM. CONTRACTION OF MUSCLE IS TRIGGERED BY
CC RELEASE OF CA++ FROM SR FOLLOWING DEPOLARIZATION OF T-TUBULES (BY
CC SIMILARITY).
CC -!- FUNCTION: INTRACELLULAR CALCIUM CHANNEL MAY BE INVOLVED IN
CC MORPHOGENETIC MOVEMENTS DURING EMBRYONIC DEVELOPMENT.
CC -!- SUBUNIT: HOMOTETRAMER (POTENTIAL).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -!- ALTERNATIVE PRODUCTS: 5 ISOFORMS; 1 (SHOWN HERE), 2, 3, 4 AND 5;
CC ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -!- DEVELOPMENTAL STAGE: ABUNDANT IN 6-12 HOUR EMBRYOS, REDUCED
CC EXPRESSION IN SECOND AND THIRD INSTAR LARVAL STAGES.
CC -!- MISCELLANEOUS: THE CALCIUM RELEASE CHANNEL IS MODULATED BY CA++,
CC MG++, ATP, AND CALMODULIN (BY SIMILARITY).
CC -!- MISCELLANEOUS: THE CALCIUM RELEASE CHANNEL ACTIVITY RESIDES IN THE
CC C-TERMINAL REGION WITH THE REMAINING PART OF THE PROTEIN
CC CONSTITUTES THE 'FOOT' STRUCTURE SPANNING THE JUNCTIONAL GAP
CC BETWEEN THE SR AND THE T-TUBULE. IT IS POSSIBLE THAT THE FOOT
CC STRUCTURE INTERACTS WITH THE CYTOPLASMIC REGION OF THE
CC DIHYDROPYRIDINE RECEPTOR (BY SIMILARITY).
CC -!- MISCELLANEOUS: RYANODINE IS AN ALKALOID THAT BINDS TO THE CA-
CC RELEASE CHANNEL IN JUNCTIONAL SR AND MODULATES ITS ACTIVITY (BY
CC SIMILARITY).
CC -!- SIMILARITY: LOCAL & LOW WITH THE NICOTINIC ACETYLCHOLINE RECEPTOR
CC (N-ACHR) SUBUNITS.
CC EMBL; D17389; BAA04212.1; ALT INIT.
CC EMBL; D17389; BAA41469.1; ALT INIT.
CC EMBL; D17389; BAA41470.1; ALT INIT.
CC EMBL; D17389; BAA41471.1; ALT INIT.
CC EMBL; AE003835; AAF59036.1; -
CC FlyBase; FBgn011286; Rva-r44F.
CC InterPro; IPR000699; Ca-rel_channel.
CC InterPro; IPR001682; Ca/Na_pore.
CC InterPro; IPR002048; EF-hand.
CC InterPro; IPR005821; Ion_trans.
CC InterPro; IPR003608; MIR.
CC InterPro; IPR001215; Ryanodn_receptor.
CC InterPro; IPR003032; RYR.
CC InterPro; IPR003877; SPRY_receptor.
CC Pfam; PF00520; Ion_trans; 1.
CC Pfam; PF02815; MIR; 4.
CC Pfam; PF01365; RYDR_ITPR; 2.
CC Pfam; PF02026; RYR; 4.
CC Pfam; PF00622; SPRY; 3.
CC PRINTS; PR00795; RYANODINER.
CC ProDom; PD000012; EF-hand; 1.
CC SMART; SM00472; MIR; 4.
CC SMART; SM00449; SPRY; 3.
CC Receptor; Transmembrane; Ionic channel; Calcium channel; Repeat;
CC Phosphorylation; Alternative splicing; Developmental protein.
FT DOMAIN 1 3207
FT TRANSMEM 3208 3228 M' (POTENTIAL).
FT TRANSMEM 3281 3299 M' (POTENTIAL).

FT TRANSMEM 4095 4114 M1 (POTENTIAL).
FT TRANSMEM 4133 4151 M2 (POTENTIAL).
FT TRANSMEM 4382 4400 M3 (POTENTIAL).
FT TRANSMEM 4467 4485 M4 (POTENTIAL).
FT TRANSMEM 4625 4646 M5 (POTENTIAL).
FT TRANSMEM 4719 4742 M6 (POTENTIAL).
FT TRANSMEM 4854 4885 M7 (POTENTIAL).
FT TRANSMEM 4902 4921 M8 (POTENTIAL).
FT TRANSMEM 4943 4962 M9 (POTENTIAL).
FT TRANSMEM 4987 5001 M10 (POTENTIAL).
FT DOMAIN 1954 1961 GLU-RICH (ACIDIC).
FT DOMAIN 825 836 6 X APPROXIMATE REPEATS.
FT REPEAT 825 936 1.
FT REPEAT 938 1055 2.
FT REPEAT 1382 1397 3 (INCOMPLETE).
FT REPEAT 1409 1444 4 (INCOMPLETE).
FT REPEAT 2800 2920 5.
FT REPEAT 2921 3043 6.
FT MOD_RES 2918 2918 PHOSPHORYLATION (BY CAPK AND CGPK)
FT MOD_RES 4428 4428 (BY SIMILARITY).
FT MOD_RES 1025 1059 PHOSPHORYLATION (POTENTIAL).
FT VARSPLIC 1115 1145 VRTLLVGYLDPTGEGTEALLAQRKLKFAGR -> SA
FT VARSPLIC 1856 1869 NAPGLRICGSSDGRNGTSGRTTQVRRIIP (IN
FT VARSPLIC 2294 2297 ISOFORM 2, ISOFORM 3, ISOFORM 4 AND
FT VARSPLIC 2262 2262 ISOFORM 5).
FT VARSPLIC 2530 2530 VTKMHAGSIHFVYRAGDVIGCFIDVKEO -> EEKVYV
FT VARSPLIC 2907 2907 GVSEFGKQCGPGIVGVFLDLADH (IN ISOFORM 3
FT VARSPLIC 3194 3194 AND ISOFORM 4).
FT VARSPLIC 3341 3341 MISSING (IN ISOFORM 4 AND ISOFORM 5).
FT VARSPLIC 4076 4076 CLREQ -> SWQ (IN REF. 1).
FT VARSPLIC 5107 AA; 578839 MW; 2AC48FDCC058F36F CRC64;
SQ SEQUENCE 5107 AA; 578839 MW; 2AC48FDCC058F36F CRC64;
Alignment Scores:
Pred. No.: 4.43 Length: 5107
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 5 Gaps: 0
US-09-774-490-1 (1-2709) x Q24498 (1-5107)
Qy 67 CGTGGGTGACACCCCTCGACGTGGCA 41
Db 2336 ArgGlySerThrProLeuAspValala 2344
RESULT 44
Q8MK55
ID Q8MK55 PRELIMINARY; PRT; 5113 AA.
AC Q8MK55
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG10844-PB.
GN RYA-R44F OR CG10844.
OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; *Drosophila*.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Randon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
Abriel J.F., Aggadani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,
Baillet R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,
de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I.I., Dietz S.M.,
Dodson K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
Durbin K.J., Gabrielson A.E., Garg N.S., Gelbart W.M., Glasser K.,
Foster C., Gabrielson A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacleb J.M.,
Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
"The genome sequence of *Drosophila melanogaster*.";
Science 287:2185-2195(2000).
[2]
SEQUENCE FROM N.A.
RA Celnik S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
Ferriera S., Frise E., Gallie R.F., Garg N.S., George R.A.,
Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunco J.,
Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B., Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
"Sequencing of *Drosophila melanogaster* genome".;
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
Tupy J.L., Bergman C., Berman B., Carlson J.W., Celnik S.E.,
Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
"Annotation of *Drosophila melanogaster* genome".;
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[4]
SEQUENCE FROM N.A.
RA Adams M.D., Celnik S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[5]
SEQUENCE FROM N.A.
RA FlyBase;
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR ENBL; AB003835; AAM71082.1; -
DR FlyBase; FBgn0011286; Rva-t44f.

DR	InterPro; IPR000699; Ca-rel_channel.
DR	InterPro; IPR001682; CaNa_pore.
DR	InterPro; IPR002048; EF-hand.
DR	InterPro; IPR005821; Ion trans.
DR	InterPro; IPR003608; MIR_
DR	InterPro; IPR001215; Ryanodn_receptor.
DR	InterPro; IPR003032; Ryr.
DR	InterPro; IPR003877; SPRY_receptor.
DR	Pfam; PF00520; ion_trans; 1.
DR	Pfam; PF02815; MIR_4.
DR	Pfam; PF01365; RYDR_ITPR; 2.
DR	Pfam; PF02026; Ryr; 4.
DR	Pfam; PF00622; SPRY; 3.
DR	PRINTS; PR00795; RYANODINER.
DR	PRODOM; PD000012; EF-hand; 1.
DR	SMART; SMO0472; MIR; 4.
DR	SMART; SMO0449; SPRY; 3.
KW	Ionic channel; Transmembrane.
SQ	SEQUENCE 5113 AA; 579656 MW; CFE36D40BE6729B5 CRC64;
 Alignment Scores:	
Pred. No.:	4.42 Length: 5113
Score:	9.00 Matches: 9
Percent Similarity:	100.00% Conservative: 0
Best Local Similarity:	100.00% Mismatches: 0
Query Match:	1.03% Indels: 0
DB:	5 Gaps: 0
 US-09-774-490-1 (1-2709) x Q8MK55 (1-5113)	
Qy	67 CGTGGGTGCACACCCCTCGACGTTGCCA 41
Db	2342 ArgGlySerThr-ProLeuAspValAla 2350
RESULT 45	
Q8MK53	
ID	PRELIMINARY; PRT; 5113 AA.
AC	Q8MK53;
DT	01-OCT-2002 (TEMBLrel. 22, Created)
DT	01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DT	01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE	CGI0844-PD.
GN	RYA-R44F OR CGI0844.
OS	Drosophila melanogaster (fruit fly).
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC	Ephydroidea; Drosophilidae; Drosophila.
OX	NCBI_TaxID=7227;
RN	[1]
RC	SEQUENCE FROM N.A.
RP	STRAIN=Berkely;
EX	MEDLINE=20196006; PubMed=10731132;
RA	Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA	Ananatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA	Abriel J.P., Agbayani A., An H.-J., Andrews-Frannkoch C., Baldwin D.,
RA	Ballew R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
RA	Beeson K.Y., Benos P.V., Beriman B.P., Bhandari D., Bolshakov S.,
RA	Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA	Cherry J.M., Clawley S., Dahlike C., Davenport L.B., Davies P.,
RA	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.
RA	Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.
RA	Fosler C., Gabrieliian A.E., Garc N.S., Gelbart W.M., Glasser K.,
RA	Glocke A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.
RA	Kimmel B.E., Kodica C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulyov G., Milshina N.V., Mobarly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamon I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celnik S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Benson J., An H., Baldwin D., Banson J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreria S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson K.A., Nunoo J.,
RA Pacleb J., Parasag V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celnik S.E.,
RA Clump M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celnik S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003835; AAM71084.1; -;
DR FlyBase; FBgn011286; Rva-r44F.
DR InterPro; IPR000659; Ca-rel channel.
DR InterPro; IPR001582; Ca/Na_Pore.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR005821; Ion trans.
DR InterPro; IPR003608; MIR.
DR InterPro; IPR001215; Ryanodn_receptor.
DR InterPro; IPR003032; RyR.
DR InterPro; IPR003877; SPRY_receptor.
DR Pfam; PF00520; ion_trans; 1.
DR Pfam; PF02815; MIR; 4.
DR Pfam; PF01365; RYDR_ITER; 2.
DR Pfam; PF02026; RyR; 4.
DR Pfam; PF00622; SPRY; 3.
DR PRINTS; PR00795; RYANODINER.
DR ProDom; PD000012; EF-hand; 1.
DR SMART; SM00472; MIR; 4.
DR SMART; SM00449; SPRY; 3.
RW Ionic channel; Transmembrane.

SQ SEQUENCE 5113 AA; 579473 MW; 6BE70740475694AB CRC64;
Alignment Scores: 4.42 Length: 5113
Pred. No.: 9.00 Matches: 9
Score: 100.00% Conservativity: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 5 Gaps: 0
US-09-774-490-1 (1-2709) x Q8MK53 (1-5113)
QY 67 CGTGGTCCGACACCCCTCGACGTGGCA 41
|||||
DB 2342 ArgglyserthrProleuaspValala 2350
RESULT 46
Q8MK54
ID Q8MK54 PRELIMINARY; PRT; 5127 AA.
AC Q8MK54;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE CGI0844-PC.
GN RYA-R44F OR CGI0844.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreria S., Fleischmann W.,
RA Fosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulyov G., Milshina N.V., Mobarly C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamon I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [2]

RP SEQUENCE FROM N.A.
 RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
 RA Bantz J., An H., Baldwin D., Banton J., Beeson K.Y., Buesam D.A.,
 RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
 RA Ferreria S., Frise E., Galle R.F., Garg N.S., George R.A.,
 RA Gonzalez M., Houch J., Hoskins R.A., Hostin D., Howland T.J.,
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
 RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
 RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
 RA Stapleton M., Strong R., Svirskaas R., Tector C., Tyler D.,
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
 RT "Sequencing of Drosophila melanogaster genome."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Mirra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
 RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
 RA Tupy J.L., Bergman C., Bernan B., Carlson J.W., Celniker S.E.,
 RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
 RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
 RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
 RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
 RT "Annotation of Drosophila melanogaster genome."
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
 RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003835; AAM71083.1; -;
 DR FlyBase; FBgn0011286; Rya-r44F.
 DR InterPro; IPR000699; Ca-rel channel.
 DR InterPro; IPR001682; Ca/Na pore.
 DR InterPro; IPR002048; EF-hand.
 DR InterPro; IPR005821; Ion trans.
 DR InterPro; IPR003608; MIR.
 DR InterPro; IPR001215; Ryanodn receptor.
 DR InterPro; IPR003032; RYR.
 DR InterPro; IPR003877; SPRY receptor.
 DR Pfam; PF00520; ion trans; 1.
 DR Pfam; PF02815; MIR; 4.
 DR Pfam; PF01365; RYDR_ITPR; 2.
 DR Pfam; PF02026; RYR; 4.
 DR Pfam; PF00622; SPRY; 3.
 DR PRINTS; PR00795; RYANODINER.
 DR ProDom; PD000012; EF-hand; 1.
 DR SMART; SM00472; MIR; 4.
 DR SMART; SM00449; SPRY; 3.
 KW Ionic channel; transmembrane.
 SQ SEQUENCE 5127 AA; 580814 MW; C4F4F3B75DCECOAC CRC64;

Alignment Scores:
 Pred. No.: 4.42 Length: 5127
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.03% Indels: 0
 DB: 5 Gaps: 0

US-09-774-490-1 (1-2709) x Q8KMS4 (1-5127)

Qy 67 CGTGGTCGACACCCCTCGAGTGCCA 41
 Db 2356 ArgGlySerThrProLeuAspValAla 2364
 RESULT 47
 Q8K4M4

ID Q8K4M4 PRELIMINARY; PRT; 48 AA.
 AC Q8K4M4;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Senaphorin 3D (Fragment).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Sv;
 RA Raper J.A., Renzi M.J., Feiner L.;
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF164784; AAM96212.1; -;
 DR InterPro; IPR001627; Sema.
 DR Pfam; PF01403; Sema; 1.
 FT NON TER 1 48
 FT NON TER 48 48
 SQ SEQUENCE 48 AA; 5415 MW; D2994B4F10F067C0 CRC64;
 Alignment Scores:
 Pred. No.: 108 Length: 48
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 11 Gaps: 0
 US-09-774-490-1 (1-2709) x Q8K4M4 (1-48)
 Qy 1199 ATTTCAAGGATCAGCGGTGTGT 1222
 Db 2 IlePheLysGlySerAlaValCys 9
 RESULT 48
 Q924C2
 ID Q924C2 PRELIMINARY; PRT; 51 AA.
 AC Q924C2;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Post-segregation killing protein.
 GN PNDA.
 OS Shigella sonnei.
 OG Plasmid Colib-P9.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Shigella.
 OX NCBI_TaxID=624;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=P9;
 RA Sempel G., Mizobuchi K.;
 RT "Organization and diversification of plasmid genomes: complete
 RT nucleotide sequence of the Colib-p9 genome."
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB021078; BAA75145.1; -;
 DR InterPro; IPR000021; Hok/gef_toxin.
 DR Pfam; PF01848; HOK_GEF; 1.
 DR PRINTS; PR00281; HOKGEFTOXIC.
 DR PROSITE; PS00556; HOK_GEF; 1.
 KW Plasmid.
 SQ SEQUENCE 51 AA; 5948 MW; 2DA97D90644FEFA9 CRC64;
 Alignment Scores:
 Pred. No.: 107 Length: 51
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x Q9Z4C2 (1-51)

QY 1374 CAAGGACCTTCCTGATGATGTTA 1397
DB 4 GlnArgThrPheLeuMetMetLeu 11

RESULT 49

OS9335 PRELIMINARY; PRT; 101 AA.
AC O59335; 07, Created)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Hypothetical protein PH1667.
GN PH1667.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=98344137; PubMed=9679194;
RA Kawabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,
RA Funahashi T., Tanaka T., Kudo Y., Yamazaki J., Kushida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
DR EMBL: AP000006; BAA30779.1; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 101 AA; 11519 MW; 846ED3C0C2CF8E7 CRC64;

Alignment Scores:

Pred. No.: 97.1 Length: 101
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 17 Gaps: 0

US-09-774-490-1 (1-2709) x O59335 (1-101)

QY 67 CTGGTGCACACCTTCGACGTG 44
DB 24 ArgGlySerThrProLeuAspVal 31

RESULT 50

OS9335 PRELIMINARY; PRT; 178 AA.
AC O8ZE20;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TRENBLrel. 20, Last annotation update)
DE Hypothetical protein YPO2006.
GN YPO2006.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebaihia M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdano-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Feltwell T., Hamlin N., Hollroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,

RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
DR EMBL: AJ414151; CAC90819.1; --
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 178 AA; 19006 MW; 86757210C29B6C6C CRC64;

Alignment Scores:
Pred. No.: 89.7 Length: 178
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x Q8ZE20 (1-178)

QY 2517 GCTCAGACACTCTGGTGGCCCTC 2494
DB 138 AlaGlnThrLeuLeuGlyAlaLeu 145

RESULT 51

OS9357 PRELIMINARY; PRT; 180 AA.
AC O9S7S7;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Flif8.39 protein.
GN Flif8.39 OR F8A24.15.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.,
RA Renning C.M., Koo H., Fujii C.Y., Utterback T.R., Barnstead M.E.,
RA Bowman C.L., White O., Niernan W.C., Fraser C.M.;
RT "Arabidopsis thaliana chromosome I BAC Flif8 genomic sequence.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC016661; AAF23308.1; --
DR EMBL: AC015985; AAF23255.1; --
DR InterPro: IPR000804; Clat_adaptor_s.
DR Pfam: PF01217; Clat_adaptor_s; 1.
KW Coat protein.
SQ SEQUENCE 180 AA; 19727 MW; 51DB46E11C67B534 CRC64;

Alignment Scores:
Pred. No.: 89.6 Length: 180
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 10 Gaps: 0

US-09-774-490-1 (1-2709) x Q9S7S7 (1-180)

QY 336 ATCATGTTGGATTCACACATT 313
DB 59 IleThrLeuLeuAspSerAsnIle 66

RESULT 52

OS9MB00 PRELIMINARY; PRT; 184 AA.
AC O9MB00;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE Nonclathrin coat protein zetal-COP.
GN COP21.

```

OS Brassica campestris (Field mustard).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OX eurosids II; Brassicales; Brassicaceae; Brassica.
RN NCBI_TaxID=3711;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=pekinensis;
RA Hahn Y., Chung J.H.;
RT "Identification of zeta-COP genes from various organisms.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL; AB040039; BAA92778.1; -.
DR InterPro; IPR000504; Clat_adaptor_s.
DR Pfam; PF01217; Clat_adaptor_s; 1.
KW Coat protein.
SQ SEQUENCE 184 AA; 20344 MW; 97C83032C008E323 CRC64;

Alignment Scores:
Pred. No.: 89.3 Length: 184
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 10 Gaps: 0

US-09-774-490-1 (1-2709) x Q9MB00 (1-184)
Qy 336 ATCACATTGTTGGATTCCCAACATT 313
Db 64 ILeThrLeuLeuAspSerAsnIle 71

RESULT 53
Q8CL36 PRELIMINARY; PRT; 218 AA.
AC Q8CL36;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DE Hypothetical.
GN Y2302.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=2137863; PubMed=12142430;
RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of Yersinia pestis KIM.";
RL J. Bacteriol. 184:4601-4611 (2002).
RL EMBL; AB013832; AAM85861.1; -.
KW Hypothetical protein.
SQ SEQUENCE 218 AA; 23531 MW; 17E162EB7B28FDB8 CRC64;

Alignment Scores:
Pred. No.: 87.2 Length: 218
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x Q8CL36 (1-218)
Qy 2517 GCTCAGACACTCTGGGTGCCCTC 2494
Db 178 AlaGlnThrLeuLeuGlyAlaLeu 185

```

```

RESULT 54
Q9KAM9 PRELIMINARY; PRT; 223 AA.
AC Q9KAM9;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Butyrate-acetoacetate CoA-transferase.
GN BH2258.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirana C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331 (2000).
DR EMBL; AP001514; BAB05977.1; -.
DR InterPro; IPR004165; CoA_trans.
DR Pfam; PF01144; CoA_trans; 1.
KW Transferase; Complete proteome.
SQ SEQUENCE 223 AA; 24448 MW; 2CCE9F28C6FB2B26 CRC64;

Alignment Scores:
Pred. No.: 86.9 Length: 223
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x Q9KAM9 (1-223)
Qy 2120 GTTGTAGACTACGTAGCAGAGGC 2097
Db 36 ValValAspTyrValAlaGluGly 43

RESULT 55
Q8HM61 PRELIMINARY; PRT; 227 AA.
AC Q8HM61;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE ATPase subunit 6.
GN ATPase 6.
OS Lophius americanus (American goosefish) (Anglerfish).
OC Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Paracanthopterygii; Lophiiformes; Lophiidae; Lophius.
OX NCBI_TaxID=8073;
[1]
RP SEQUENCE FROM N.A.
RA Miya M.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Miya M., Takeshima H., Endo H., Ishiguro N.B., Inoue J.G., Mukai T.,
RA Satoh T.P., Yamaguchi M., Kawaguchi A., Mabuchi K., Shirai S.M.,
RA Nishida M.;
RT "Major Patterns of higher teleostean phylogenies: A new perspective
RT based on 100 complete mitochondrial DNA sequences.";
RL Mol. Phylogenet. Evol. 26:121-138 (2002).
DR EMBL; AP004414; BAC23299.1; -.
KW Mitochondrion.
SQ SEQUENCE 227 AA; 25102 MW; 4312A5A602AD4DBD CRC64;

```

Alignment Scores:
 Pred. No.: 86.7 Length: 227
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 8 Gaps: 0

US-09-774-490-1 (1-2709) x Q8HM61 (1-227)

QY 1248 CTTCTCACATCACTCATGCTATAC 1225
 |||||
 Db 72 LeuLeuThrSerLeuMetLeuTyr 79

RESULT 56
 Q96ID9 PRELIMINARY; PRT; 277 AA.

ID Q96ID9
 AC Q96ID9
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
 DE Hypothetical protein.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=eye;
 RA Strausberg R.;
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
 DR EMBL: BC007602; AA07602.1; --
 DR InterPro; IPR002350; kazal.
 DR InterPro; IPR001841; ZnF_ring.
 DR Pfam; PF00097; zf-C3HC4; 1.
 DR SMART; SM00184; RING; 1.
 DR PROSITE; PS00282; KAZAL; 1.
 DR PROSITE; PS00518; ZF_RING_1; 1.
 DR PROSITE; PS00089; ZF_RING_2; 1.
 KW Hypothetical protein; Metal-binding; Zinc; Zinc-finger.
 SQ SEQUENCE 277 AA; 30443 MW; EBED1BD2B7B243F0 CRC64;

Alignment Scores:
 Pred. No.: 84.3 Length: 277
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q96ID9 (1-277)

QY 1801 GTAAGGTCCTCGGCGAGCAACA 1778
 |||||
 Db 204 ValArgValSerGlyGluAlaThr 211

RESULT 57
 Q9V3F8 PRELIMINARY; PRT; 280 AA.

ID Q9V3F8
 AC Q9V3F8
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE P5CR protein.
 GN P5CR OR CG6009.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=107311132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclab J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=CS;
 RA Misener S.R., Walker V.K.;
 RT "High density of unrelated genes showing overlapping and intratronic
 RT transcription units in Drosophila."
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003724; AAF55626.1; --
 DR EMBL; AF170829; AAD49740.1; --
 DR FlyBase; FBgn0015781; P5CR.
 DR InterPro; IPR000304; P5CR.
 DR Pfam; PF01089; P5CR; 1.
 DR TIGRFAMs; TIGR00112; proC; 1.
 DR PROSITE; PS00521; P5CR; 1.
 KW Oxidoreductase.
 SQ SEQUENCE 280 AA; 29621 MW; D3D4FD9611574684 CRC64;

Alignment Scores:
 Pred. No.: 84.2 Length: 280
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 5 Gaps: 0

US-09-774-490-1 (1-2709) x Q9V3F8 (1-280)

QY 2520 GCAGCTCAGACACTCTCGGTGCC 2497
 |||||
 Db 215 AlaAlaGlnThrLeuLeuGlyAla 222

RESULT 58
 O96643

ID O96643 PRELIMINARY; PRT; 280 AA.
AC Q96643;
DT 01-MAY-1999 (TReMBLrel. 10, Created)
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Pyroline 5-carboxylate reductase.
GN P5CR OR C6609.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CS;
RA Misener S.R., Walker V.K.;
RT "Cloning and sequence analysis of the gene encoding pyroline 5-
RT carboxylate reductase from Drosophila melanogaster.";
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF098020; AAC70780.1; -.
DR FlyBase; FBgn0015781; P5CR.
DR InterPro; IPR00304; P5CR.
DR Pfam; PF01089; P5CR; 1.
DR TIGRFAMs; TIGR00112; proC; 1.
DR PROSITE; PS00521; P5CR; 1.
SQ SEQUENCE 280 AA; 29688 MW; 2A7416D39D2A2369 CRC64;
Alignment Scores:
Pred. No.: 84.2 Length: 280
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 5 Gaps: 0
US-09-774-490-1 (1-2709) x O96643 (1-280)
QY 2520 GCAGCTCAGACTCTCGGTGCC 2497
Db 215 AlaAlaGlnThrLeuLeuGlyAla 222
RESULT 59
O22841
ID O22841 PRELIMINARY; PRT; 283 AA.
AC O22841;
DT 01-JAN-1998 (TReMBLrel. 05, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
DE Putative endochitinase.
GN AT2G43620.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.B., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC002333; AAB64044.1; -.
DR HSSP; P02876; 9WGA.
DR InterPro; IPR001002; Chitin binding_1.
DR InterPro; IPR000726; Glyco_hydro_19.
DR Pfam; PF00187; Chitin_bind_1; 1.
DR Pfam; PF00182; Glyco_hydro_19; 1.
DR PRINTS; PR00451; CHITINBINDNG.
DR ProDom; PD000609; Chitin binding_1; 1.
DR SMART; SM00270; ChtBD1; 1.
DR PROSITE; PS00026; CHITIN_BINDING; 1.
KW Chitin-binding.
SQ SEQUENCE 283 AA; 30377 MW; D758222971EB8404 CRC64;
Alignment Scores:
Pred. No.: 84.1 Length: 283
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 10 Gaps: 0
US-09-774-490-1 (1-2709) x O22841 (1-283)
QY 2319 TGCAGCTCATCAACCAACCCCAATC 2342
Db 64 CysSerSerThrThrProile 71
RESULT 60
Q91632
ID Q91632 PRELIMINARY; PRT; 292 AA.
AC Q91632;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE BTEE.
GN GENE 1.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96312514; PubMed=8700860;
RA Brown D.D., Wang Z., Furlow J.D., Kanamori A., Schwartzman R.A.,
RA Remo B.F., Pinder A.;
RT "The thyroid hormone-induced tail resorption program during Xenopus
RT laevis metamorphosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:1924-1929(1996).
DR EMBL; U35408; AAC59863.1; -.
DR HSSP; P08047; 1SP2
DR InterPro; IPR007087; Znf_C2H2.
DR ProDom; PD000096; zf_C2H2; 3.
DR Pfam; PF00003; Znf_C2H2; 1.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 3.
KW Metal-binding; zinc; Zinc-finger.
FT VARIANT 16 16 V -> M.
SQ SEQUENCE 292 AA; 32789 MW; EF62BD68DD9CA2C3 CRC64;
Alignment Scores:
Pred. No.: 83.7 Length: 292
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 13 Gaps: 0
US-09-774-490-1 (1-2709) x Q91632 (1-292)
QY 2316 TCATGCAGCTCATCAACCAACCCCA 2339

```
Db 140 SerCysSerSerThrPro 147
|||||
RESULT 61
Q91633 PRELIMINARY; PRT; 292 AA.
AC Q91633;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE BTEB.
GN XENOPUS laevis (African clawed frog).
OS XENOPUS laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96312514; PubMed=8700860;
RA Brown D.D., Wang Z., Furlow J.D., Kanamori A., Schwartzman R.A.,
RA Remo B.F., Pinder A.;
RT "The thyroid hormone-induced tail resorption program during Xenopus
RT laevis metamorphosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:1924-1929(1996).
DR EMBL; U35409; AAC59864.1; -.
DR HSSP; P08047; 1SP2.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00096; zf-C2H2; 3.
DR ProDom; PD000003; Znf_C2H2; 1.
DR SMART; SM00355; Znf_C2H2; 3.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 3.
DR PROSITE; PS00157; ZINC_FINGER_C2H2_2; 3.
KW Metal-binding; Zinc; Zinc-finger.
SQ SEQUENCE 292 AA; 32748 MW; 812BA9D57EFBCFA CRC64;

Alignment Scores:
Pred. No.: 83.7 Length: 292
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 13 Gaps: 0

US-09-774-490-1 (1-2709) x Q91633 (1-292)

QY 2316 TCATCAGCTCATCAACCAACCCCA 2339
|||||
Db 140 SerCysSerSerThrPro 147

RESULT 62
Q9D3M7 PRELIMINARY; PRT; 294 AA.
AC Q9D3M7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Arac-type DNA-binding domain-containing protein.
GN V21661.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016813; AAC08519.1; -.
KW DNA-binding; Complete proteome.
SQ SEQUENCE 294 AA; 33902 MW; A62C99A081BE20E2 CRC64;
```

```
Alignment Scores:
Pred. No.: 83.6 Length: 294
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x Q9D3M7 (1-294)

QY 991 TCTAGCGTAGCTAGCTTTCCAGA 968
|||||
Db 91 SerSerValSerSerPheSerArg 98

RESULT 63
Q9D3Q3 PRELIMINARY; PRT; 299 AA.
AC Q9D3Q3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Spermidine synthase.
GN V21635.
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016813; AAC08493.1; -.
KW Complete proteome.
SQ SEQUENCE 299 AA; 34010 MW; 31687212D98C7C6A CRC64;

Alignment Scores:
Pred. No.: 83.4 Length: 299
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x Q9D3Q3 (1-299)

QY 9 ATTTATCGATGTTAACAAGCTTA 32
|||||
Db 237 IleuSerMetLeuThrSerLeu 244

RESULT 64
Q9A6S1 PRELIMINARY; PRT; 309 AA.
AC Q9A6S1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein CC2012.
GN CC2012.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacterales;
OC Caulobacteraceae; Caulobacter.
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 19089 / CB15;
RA MEDLINE=21173698; PubMed=11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
```

```
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
RA Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,  
RT "Complete genome sequence of Caulobacter crescentus";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
DR EMBL; AE005874; AAK23987.1; -.  
DR TIGR; CC2012; -.  
DR InterPro; IPR002052; N6_Mtase.  
DR InterPro; IPR002965; P_Rich_extensn.  
DR PRINTS; PR01217; PRICHEXTEN.  
DR PROSITE; PS00092; N6_MTASE; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 309 AA; 34342 MW; 73FC082AE5D8E6A8 CRC64;  
  
Alignment Scores:  
Pred. No.: 83 Length: 309  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.92% Indels: 0  
DB: 16 Gaps: 0  
  
US-09-774-490-1 (1-2709) x Q9A6S1 (1-309)  
  
Qy 1074 ATCAGACGAGCTTTGAGGAAGTT 1051  
Db 38 IileArgAlaLeuArgAsnVal 45  
  
RESULT 65  
ID Q8NH68 PRELIMINARY; PRT; 312 AA.  
AC Q8NH68;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Seven transmembrane helix receptor.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,  
RA Tautsuni S., Aburatani H., Asai K., Akiyama Y.;  
RT "Genome-wide discovery and analysis of human seven transmembrane helix  
receptor genes";  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBDJ databases.  
DR EMBL; AB065521; BAC05769.1; -.  
DR InterPro; IPR000276; GPCR_Rhodpsn.  
DR Pfam; PF00001; 7tm.1; 1.  
DR PRINTS; PR00237; GPCRHHODOPSN.  
DR PROSITE; PS00237; G_PROTEIN_REC_F1_1; 1.  
DR PROSITE; PS00262; G_PROTEIN_REC_F1_2; 1.  
KW Receptor; Transmembrane.  
SQ SEQUENCE 312 AA; 34897 MW; 851854030D2BEE6A CRC64;  
  
Alignment Scores:  
Pred. No.: 82.9 Length: 312  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.89% Indels: 0  
DB: 4 Gaps: 0  
  
US-09-774-490-1 (1-2709) x Q8NH68 (1-312)  
  
Qy 1770 GTGCTGAGTGTGCTCGCCGAG 1793  
Db 225 ValLeuSerValAlaSerProGlu 232  
  
RESULT 66  
Q97D40
```

```
ID Q97D40 PRELIMINARY; PRT; 326 AA.  
AC Q97D40;  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)  
DE Oligopeptide ABC transporter, ATPase component.  
GN CAC3641.  
OS Clostridium acetobutylicum.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI_TaxID=1488;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=ATCC 824 / DSM 792 / VKM B-1787;  
RX MEDLINE=21359325; PubMed=11466286;  
RA Noelling J., Breston G., Omelchenko M.V., Makarova K.S., Zeng Q.,  
RA Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,  
RA Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,  
RA Bennett G.N., Koonin E.V., Smith D.R.;  
RT "Genome sequence and comparative analysis of the solvent-producing  
bacterium Clostridium acetobutylicum";  
RL J. Bacteriol. 183:4823-4838(2001).  
DR EMBL; AE007859; AAK81563.1; -.  
DR InterPro; IPR003593; AAA_ATPase.  
DR InterPro; IPR003439; ABC_transporter.  
DR Pfam; PF00005; ABC_tran; 1.  
DR ProDom; PD000006; ABC_transporter; 1.  
DR SMART; SM00382; AAA; 1.  
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.  
KW Complete proteome.  
SQ SEQUENCE 326 AA; 36325 MW; 9F1A2DE19ACAD435 CRC64;  
  
Alignment Scores:  
Pred. No.: 82.4 Length: 326  
Score: 8.00 Matches: 8  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.89% Indels: 0  
DB: 16 Gaps: 0  
  
US-09-774-490-1 (1-2709) x Q97D40 (1-326)  
  
Qy 296 AAATTATCTACAAAGAAATGTTG 319  
Db 84 LysLeuSerTyrIysGluMetLeu 91  
  
RESULT 67  
ID Q91914 PRELIMINARY; PRT; 327 AA.  
AC Q91914;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)  
DE Homeobox protein.  
GN XOM.  
OS Xenopus laevis (African clawed frog);  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96312918; PubMed=8756284;  
RA Lader R.K., Mohun T., Smith C., Snape A.M.;  
RT "Xom: a Xenopus homeobox gene which mediates the early effects of BMP-  
4";  
RL Development 122:2385-2394(1996).  
CC -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
DR EMBL; X98454; CAA67093.1; -.  
DR HSP; P02836; IENH.  
DR InterPro; IPR001356; Homeobox.  
DR Pfam; PF00046; homeobox; 1.  
DR PRINTS; PR00024; HOMEBOX.
```

DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX_1; 1.
DR PROSITE; PS00071; HOMEBOX_2; 1.
DR DNA-binding; Homeobox; Nuclear protein.
SQ SEQUENCE 327 AA; 36675 MW; AB669084F355F4E4 CRC64;

Alignment Scores:
Pred. No.: 82.4 Length: 327
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 13 Gaps: 0

US-09-774-490-1 (1-2709) x Q91914 (1-327)

QY 2310 GAGACTTCATGCAGCTCATCAACC 2333
Db 145 GluThrSerCysSerSerSerThr 152

RESULT 68

Q97D45 PRELIMINARY; PRT; 350 AA.
AC Q97D45;
DT 01-OCT-2001 (Tremblrel. 18, Created)
DT 01-OCT-2001 (Tremblrel. 18, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Oligopeptide ABC transporter, ATPase component.
GN CAC3635.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J.; Breton G.; Omelchenko M.V.; Makarova K.S.; Zeng Q.,
RA Gibson R.; Lee H.M.; Dubois J.; Qiu D.; Hitti J.; Wolf Y.I.,
RA Tatusov R.L.; Sabathe F.; Doucette-Stamm L.; Soucaille P.; Daly M.J.,
RA Bennett G.N.; Koonin E.V.; Smith D.R.;
RT "Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum";
RL J. Bacteriol. 183:4823-4838(2001).
DR EMBL; AE007859; AAK81558.1; -.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR003439; ABC transporter.
DR Pfam; PF00005; ABC tran; 1.
DR ProDom; PD000006; ABC transporter; 1.
DR SMART; SM00382; AAA_1
DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW Complete proteome.
SQ SEQUENCE 350 AA; 39445 MW; 33B619DF480EB51D CRC64;

Alignment Scores:
Pred. No.: 81.6 Length: 350
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x Q97D45 (1-350)

QY 296 AAATATCTCAAGAAGATGTG 319
Db 108 LysLeuSerTyLysGluMetLeu 115

RESULT 69

Q9BYE7 PRELIMINARY; PRT; 352 AA.
ID Q9BYE7
AC Q9BYE7;

DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE HMBLR protein.
GN HMBLR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Takahashi N.; Akasaka T.; Nakayama T.; Koseki H.; Koga H.;
RT "Identification and characterization of hMBLR";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Cervix;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AB047006; BAB40779.1; -.
DR EMBL; BC010235; AAH10235.1; -.
DR InterPro; IPR002350; kazal.
DR InterPro; IPR001841; znf_ring.
DR Pfam; PF00097; zf-C3HC4; 1.
DR SMART; SM00184; RING; 1.
DR PROSITE; PS00282; KAZAL; 1.
DR PROSITE; PS00518; ZF_RING_1; 1.
DR PROSITE; PS00089; ZF_RING_2; 1.
KW Metal-binding; Zinc; zinc-finger.
SQ SEQUENCE 352 AA; 39241 MW; 559F971A86D9A093 CRC64;

Alignment Scores:
Pred. No.: 81.5 Length: 352
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x Q9BYE7 (1-352)

QY 1801 GTAAGGCTCTCGGCGAGCAACA 1778
Db 279 ValArgValSerGlyGluAlaThr 286

RESULT 70

Q99NA9 PRELIMINARY; PRT; 353 AA.
ID Q99NA9
AC Q99NA9;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE HMBLR protein (RIKEN cDNA 4933407A11 gene).
GN 4933407A11RIK OR HMBLR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RA Takahashi N.; Akasaka T.; Nakayama T.; Koseki H.; Koga H.;
RT "Identification and characterization of hMBLR";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.
DR EMBL; AB047007; BAB40780.1; -.
DR EMBL; BC016195; AAH16195.1; -.

DR MGI:1918291; 4933407A11RIK.
 DR InterPro; IPR002350; kazal.
 DR InterPro; IPR001841; znf ring.
 DR Pfam; PF00097; zf-C3HC4; 1.
 DR SMART; SM00184; RING; 1.
 DR PROSITE; PS00282; KAZAL; 1.
 DR PROSITE; PS00518; ZF_RING_2; 1.
 DR PROSITE; PS00089; ZF_RING_2; 1.
 KW Metal-binding; Zinc; Zinc-finger.
 SQ SEQUENCE 353 AA; 39820 MW; 0D29B7FE065B5448 CRC64;

Alignment Scores:
 Pred. No.: 81.5 Length: 353
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x Q99NA9 (1-353)

QY 1801 CTAGGGTCTCGGCGAGCAACA 1778
 |||||
 Db 280 ValArgValSerGlyGluAlaThr 287

RESULT 71
 Q90753
 ID Q90753 PRELIMINARY; PRT; 371 AA.
 AC Q90753;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Host shut-off factor (Fragment).
 GN UL41.
 OS Feline herpesvirus (Feline herpesvirus 1).
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Alphaherpesvirinae; Varicellovirus.
 OX NCBI_TaxID=10334;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B927;
 EX MEDLINE=98141690; PubMed=9482586;
 RA Willoughby K., Bennett M., Williams R.A., McCracken C., Gaskell R.M.;
 RT "Sequences of the ribonucleotide reductase-encoding genes of feline
 herpesvirus 1 and molecular phylogenetic analysis.";
 RL Virus Genes 15:203-218 (1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B927;
 RA Willoughby K.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AJ006454; CAA07026.1; -;
 DR InterPro; IPR000513; Exo N.I.
 DR InterPro; IPR002327; Virus_HS.
 DR Pfam; PF01550; Virus_HS; 1.
 FT NON TER 1
 SQ SEQUENCE 371 AA; 42219 MW; 10672R8EFDCAF675 CRC64;

Alignment Scores:
 Pred. No.: 81 Length: 371
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.92% Indels: 0
 DB: 12 Gaps: 0

US-09-774-490-1 (1-2709) x Q90753 (1-371)

QY 2498 CCTCTCAATTCGTGGTCTCC 2475
 |||||
 Db 289 ProSerGlnIleArgGlySerSer 296

RESULT 72

Q9AQP4
 ID Q9AQP4 PRELIMINARY; PRT; 386 AA.
 AC Q9AQP4;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Putative ABC transporter subunit.
 OS Pseudomonas sp. CA10.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=135214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CA10;
 RX MEDLINE=21264379; PubMed=11371531;
 RA Nojiri H., Sekiguchi H., Maeda K., Urata M., Nakai S., Yoshida T.,
 RA Habe H., Omori T.;
 RT "Genetic characterization and evolutionary implications of car gene
 cluster in carbazole-degrader, Pseudomonas sp. strain CA10.";
 RL J. Bacteriol. 183:3663-3679 (2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CA10;
 RX MEDLINE=97386424; PubMed=9244273;
 RA Sato S., Ouchiya N., Kimura T., Nojiri H., Yamane H., Omori T.;
 RT "Cloning of genes involved in carbazole degradation of Pseudomonas sp.
 strain CA10: nucleotide sequence of genes and characterization of
 meta-cleavage enzymes and hydrolase.";
 RL J. Bacteriol. 179:4841-4849 (1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CA10;
 RX MEDLINE=97386425; PubMed=9244274;
 RA Sato S., Nam J., Kauga K., Nojiri H., Yamane H., Omori T.;
 RT "Identification and characterization of genes encoding carbazole 1,9a-
 dihydroxylase in Pseudomonas sp. strain CA10.";
 RL J. Bacteriol. 179:4850-4858 (1997).
 DR EMBL; AB047548; BAB32743.1; -;
 SQ SEQUENCE 386 AA; 41689 MW; 7B6F4D2292D46BCE CRC64;

Alignment Scores:
 Pred. No.: 80.5 Length: 386
 Score: 8.00 Matches: 8
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.89% Indels: 0
 DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x Q9AQP4 (1-386)

QY 1734 TCCCTTTACACCGGTGTATTT 1757
 |||||
 Db 92 SerLeuYrThrGlyValIlePhe 99

RESULT 73
 Q8GI39
 ID Q8GI39 PRELIMINARY; PRT; 386 AA.
 AC Q8GI39;
 DT 01-MAR-2003 (TREMBLrel. 23, Created)
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
 DE Putative ABC transporter subunit.
 OS Pseudomonas resinovorans.
 OG Plasmid pCARI.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=53412;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Maeda K., Nojiri H., Shintani M., Yoshida T., Habe H., Omori T.;
 RT "Complete nucleotide sequence of carbazole/dioxin-degrading plasmid
 pCARI in Pseudomonas resinovorans strain CA10 indicates its mosaicity
 and the presence of large catabolic transposon Tn4676.";

RL J. Mol. Biol. 0:0-0(2002).
DR EMBL; AB088420; BAC41522.1; --
KW Plasmid.
SQ SEQUENCE 386 AA; 41461 MW; C9A34BAF2DD24F7A CRC64;

Alignment Scores:
Pred. No.: 80.5 Length: 386
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x Q8GI39 (1-386)

QY 1734 TCCTTTACACGGTGTGATATT 1757

Db 92 SerLeuTyThrGlyValIlePhe 99

RESULT 74

Q8RTH6 PRELIMINARY; PRT; 402 AA.

AC Q8RTH6;

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DT 01-WAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Putative O antigen ligase.

GN WAAL.

OS Vibrio cholerae.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;

OC Vibrionaceae; Vibrio.

OX NCBI_TaxID=666;

RN [1]

RP SEQUENCE FROM N.A.

RA Nesper J., Kraiss A., Schild S., Blass J., Klose K.E., Bockemuehl J.,

RA Reidl J.;

RT "Comparative and genetic analysis of the putative Vibrio cholerae LPS

core oligosaccharide biosynthesis (waa) gene cluster";

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF449194; AAL77364.1; --

DR InterPro; IPR000504; RNA_rec_mot.

DR PROSITE; PS00030; RRM_RNP_1; 1.

KW Ligase.

SQ SEQUENCE 402 AA; 46039 MW; B44948E5DF833C0E CRC64;

Alignment Scores:

Pred. No.: 80 Length: 402
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x Q8RTH6 (1-402)

QY 761 AGTATAATCTCCATCTATATAA 738

Db 347 SerIleIleLeuHisLeuLeuLys 354

RESULT 75

Q52274

ID Q52274 PRELIMINARY; PRT; 409 AA.

AC Q52274;

DT 01-JUN-1998 (TrEMBLrel. 06, Created)

DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)

DE Specificity subunit.

GN WSDS.

OS Lactococcus lactis subsp. lactis bv. diacetylactis.

OC Plasmid pND861.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.

OX NCBI_TaxID=44688;

RN [1]

RP SEQUENCE FROM N.A.
RC STRAIN=UK19161;
RX MEDLINE=20541281; PubMed=11092734;
RA Deng Y.W., Liu C.Q., Dunn N.W.;
RT "Lidi, a plasmid-encoded type I restriction and modification system in
RT Lactococcus lactis";
RL DNA Seg. 11:239-245(2000).
DR EMBL; AF034786; AAB91417.1; --
DR InterPro; IPR000055; Rest_mod_DNA.
DR Pfam; PF01420; Methylase_S; 2.
KW Plasmid.
SQ SEQUENCE 409 AA; 46727 MW; 3B26DDA650A274BC CRC64;

Alignment Scores:

Pred. No.: 79.9 Length: 409
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x O52274 (1-409)

QY 1135 GAATACATCCTGCAGTTCATCAA 1112

Db 117 GluTyrIleLeuGlnPheIleLys 124

Search completed: August 3, 2003, 10:56:00
Job time : 268.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 3, 2003, 10:29:37 ; Search time 86.5 Seconds
(without alignments)
9941.972 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 895

Sequence: 1 aatctttttatttcgatg.....aggcttttttctctaataacc 2709

Scoring table: OLIGO
Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2062474

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09774490/runat 03082003.102927.4441/app_query.fasta_1.2887
-DB=A_Geneseq19Jun03 -QFWT=fascan -SUFFIX=oligo.rag -MINMATCH=0.1 -DOORCL=0
-DOOREXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=100 -DOALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=75 -MODE=LOCAL
-OUTFWT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09774490@cgn 1.1.114 @runat 03082003.102927.4441 -NCPU=6 -ICPU=3
-NO MWAP -LARGQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : A_Geneseq19Jun03.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	771	86.1	771	16	AAR71380	Human semaphorin I
2	771	86.1	771	22	AAG62726	Amino acid sequenc
3	771	86.1	771	23	ABG96413	Human ovarian canc
4	437	48.8	796	19	AAV21264	Human semaphorin I
5	434	48.5	477	16	AAR74175	Human collapsin.
6	151	16.9	151	22	ABG54800	Human liver peptid
7	151	16.9	151	23	ABG42930	Human peptide enco
8	81	9.1	81	19	AAV21293	Human semaphorin I
9	77	8.6	77	19	AAV21288	Human semaphorin I
10	74	8.3	74	19	AAV21291	Human semaphorin I
11	69	7.7	69	19	AAV21274	Human semaphorin I
12	64	7.2	64	19	AAV21272	Human semaphorin I
13	57	6.4	57	19	AAV21299	Human semaphorin I
14	55	6.1	85	19	AAV21271	Human semaphorin I
15	53	5.9	57	22	ABG54764	Human liver peptid
16	53	5.9	57	22	ABG39696	Peptide #7202 enco
17	53	5.9	57	22	ABG60414	Human brain expres
18	53	5.9	57	22	AAW73050	Human bone marrow
19	53	5.9	57	22	AAW33272	Peptide #7309 enco
20	53	5.9	57	23	ABG42894	Human peptide enco
21	51	5.7	51	19	AAV21280	Human semaphorin I
22	50	5.6	50	19	AAV21346	Human semaphorin I
23	48	5.4	48	19	AAV21324	Human semaphorin I
24	46	5.1	46	19	AAV21331	Human semaphorin I
25	42	4.7	42	19	ABG51322	Human semaphorin I
26	41	4.6	41	22	ABG55921	Human liver peptid
27	41	4.6	41	22	ABG61385	Human brain expres
28	41	4.6	41	23	ABG44063	Human peptide enco
29	33	3.7	33	19	AAV21266	Human semaphorin I
30	33	3.7	33	22	ABG54883	Human liver peptid
31	33	3.7	33	22	ABG39792	Peptide #7298 enco
32	33	3.7	33	22	ABG24410	Protein #6409 enco
33	33	3.7	33	22	AAW60513	Human brain expres
34	33	3.7	33	22	AAW73166	Human bone marrow
35	33	3.7	33	22	AAW33379	Peptide #7416 enco
36	33	3.7	33	23	ABG43013	Human peptide enco
37	31	3.5	31	19	AAV21292	Human semaphorin I
38	30	3.4	30	19	AAV21279	Human semaphorin I
39	27	3.0	27	19	AAV21269	Human semaphorin I
40	26	2.9	26	19	AAV21296	Human semaphorin I
41	25	2.8	25	19	AAV21328	Human semaphorin I
42	24	2.7	24	19	AAV21285	Human semaphorin I
43	23	2.6	23	19	AAV21337	Human semaphorin I
44	23	2.6	23	19	AAV21310	Human semaphorin I
45	23	2.6	23	19	AAV21290	Human semaphorin I
46	22	2.5	22	19	AAV21334	Human semaphorin I
47	22	2.5	67	21	AAW02551	Human secreted pro
48	21	2.3	21	19	AAV21342	Human semaphorin I
49	21	2.3	21	19	AAV21343	Human semaphorin I
50	21	2.3	21	19	AAV21278	Human semaphorin I
51	20	2.2	20	19	AAV21329	Human semaphorin I
52	19	2.1	19	19	AAV21318	Human semaphorin I
53	19	2.1	19	19	AAV21304	Human semaphorin I
54	18	2.0	18	19	AAV21309	Human semaphorin I
55	17	1.9	17	19	AAV21301	Human semaphorin I
56	17	1.9	17	22	AAW62702	Hanatoxin-like seq
57	17	1.9	779	22	AAB84219	Amino acid sequenc
58	17	1.9	779	23	AAW78481	Human ZSMF-16
59	17	1.9	782	21	AAB23609	Human secreted pro
60	17	1.9	782	22	AAW65620	Novel human protei
61	17	1.9	785	21	AAB23636	Human secreted pro
62	17	1.9	875	22	AAW65619	Novel human protei
63	16	1.8	17	22	AAW62704	Hanatoxin-like seq
64	15	1.7	15	19	AAV21300	Human semaphorin I
65	15	1.7	15	19	AAV21281	Human semaphorin I
66	15	1.7	749	22	AAW62727	Amino acid sequenc
67	15	1.7	751	20	AAW30617	Human semaphorin E

68 15 1.7 751 21 AAB28379 Clone BR533 4. Ho
69 15 1.7 751 22 AAG62728 Amino acid sequenc
70 15 1.7 751 23 ABP68623 Human pancreatic c
71 14 1.6 14 19 AAY21315 Human semaphorin I
72 14 1.6 14 19 AAY21276 Human semaphorin I
73 14 1.6 777 20 AAY21276 Human brain tissue
74 14 1.6 777 21 AAY99427 Human PRO1491 (UNQ
75 14 1.6 777 22 AAY29197 Human PRO polypept
76 14 1.6 777 22 AAB66176 Protein of the inv
77 14 1.6 777 24 ABU71285 Human PRO1491 prot
78 14 1.6 777 24 ABU65742 Human secreted/tra
79 14 1.6 777 24 ABU66075 Novel human secret
80 14 1.6 777 24 ABU67579 Human secreted/tra
81 14 1.6 777 24 ABU65437 Human PRO polypept
82 14 1.6 777 24 ABU58573 Human PRO polypept
83 14 1.6 777 24 ABU58109 Human secreted/tra
84 14 1.6 777 24 ABU57104 Human PRO polypept
85 14 1.6 777 24 ABU10683 Human secreted/tra
86 14 1.6 807 22 AAG62729 Amino acid sequenc
87 13 1.5 13 19 AAY21302 Human semaphorin I
88 13 1.5 17 22 AAG62707 Human semaphorin I
89 13 1.5 17 22 AAG62705 Hanatoxin-like seq
90 13 1.5 775 19 AAW63748 Human semaphorin.
91 13 1.5 775 20 ABR473090 Mouse semaphorin H
92 13 1.5 775 24 ABR47588 Breast cancer asso
93 13 1.5 777 20 AAY43091 Mouse semaphorin H
94 13 1.5 777 22 AAG62730 Amino acid sequenc
95 13 1.5 785 22 AAG62731 Amino acid sequenc
96 12 1.3 12 19 AAY21325 Human semaphorin I
97 12 1.3 12 19 AAY21336 Human semaphorin I
98 12 1.3 12 19 AAY21320 Human semaphorin I
99 12 1.3 12 19 AAY21270 Human semaphorin I
100 12 1.3 12 19 AAY21275 Human semaphorin I

ALIGNMENTS

RESULT 1
ID AAR71380
XX AAR71380 standard; Protein; 771 AA.
AC AAR71380;
XX
DT 25-MAR-2003 (updated)
DT 21-NOV-1995 (first entry)
XX
DE Human semaphorin III protein.
XX
KW Semaphorin; grasshopper; human; vaccinia virus; Drosophila; Tribolium;
KW variola major virus; smallpox; semaphorin receptor binding activity;
KW modulation; nerve cell growth; immune response; viral pathogenesis;
KW neurological disease; neuro-regeneration; oncological infection.
XX
OS Homo sapiens.
XX.
PN WO9507706-A1.
XX
PD 23-MAR-1995.
XX
PF 13-SEP-1994; 94WO-US10151.
XX
PR 13-SEP-1993; 93US-0121713.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Bentley DR, Goodman CS, Kolodkin AL, Matthes D;
PI O'Connor T;
XX
XX WPI; 1995-131177/17.
DR N-PSDB; AAQ87442.
XX
PT New class of semaphorin peptide(s) and polypeptide(s) - are
potent modulators of nerve cell growth and regeneration

XX Example 2; Page 60-63; 101pp; English.
PS The sequence of the human semaphorin III protein. The proteins
XX encoded by the grasshopper semaphorin I (AAQ87441), human semaphorin
CC III, vaccinia virus semaphorin IV (AAQ87443), Drosophila semaphorin I and
CC II (AAQ87444-5), Tribolium semaphorin I (AAQ87446) or variola major
CC (smallpox) virus semaphorin IV (AAQ87447) genes were used to generate a
CC series of peptides (AAR70370-R70418), which retain semaphorin receptor
CC binding activity. The semaphorin derived or semaphorin receptor derived
CC peptides are potent modulators of nerve cell growth, immune
CC responsiveness and viral pathogenesis. They can be used in diagnosis and
CC treatment of neurological disease and neuro-regeneration, immune
CC modulation and diagnosis and treatment of viral and oncological infection
CC and diseases.
XX (Updated on 25-MAR-2003 to correct PN field.)
SQ Sequence 771 AA;

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 16 Gaps: 0

US-09-774-490-1 (1-2709) x AAR71380 (1-771)
Qy 200 ATGGGCTGTTAACTAGGATTGCTGCTTTTGGGGAGTATTACTACAGCAGCA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
Qy 260 AACTATCAGATGGGAAGAACAATGTCACAGGCTGAANTATCTCTACAAAGAATGTTG 319
Db 21 AsnTyrGlnAsnGlyLysAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
Qy 320 GAATCCAAACAATGTGATCACTTTCAATGGCTGGCCCAACAGCTCCAGTTATCATACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
Qy 380 CTTTGGATGAGGAACGGAGTAGGCTGTATGTTGGAGCAAGGATCATATTTTCATTTC 439
Db 61 LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisLeuPheSerPhe 80
Qy 440 GACCTGTTAATATCAAGGATTTTCAAAAGATTGTGTGGCCAGTATCTTACACCAAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
Qy 500 GATGAATGCAAGTGGGCTGGAAAGACATCTGAAAGAATGTGCTAATTTTCATCAAGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
Qy 560 CTTAAGGCATATAATCAGACTCACTGTGTACGCTGTGGAAACGGGGGCTTTTCATCCAATT 619
Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProile 140
Qy 620 TGCACCTACATTGAATTGGACATCATCTCCAGGACATATTTTAACTGGGAGCACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
Qy 680 CATTTTGAAAACCGCGCTGGGAAGAGTCCATATGACCTTAAGCTGTGCAGACATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
Qy 740 TTAATAGATGAGCAATATATCTCTGGAACCTGACGCTGATTTTATGGGCGAGACTTTGCT 799
Db 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
Qy 800 ATCTTCGAACTCTTGGGCACACCCACCAATCAGGACAGCAGCATGATTCAGGCTGG 859
Db 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220

Claim 6; Page 11; 29pp; English.

PS The present sequence represents a semaphorin. The specification
XX describes hanatoxin-like sequences (HTLS) found in the semaphorin
CC domain of mammalian secreted semaphorins. Hanatoxin is a tarantula
CC toxin that selectively blocks some voltage-gated potassium and calcium
CC channels. The HTLS is responsible for the dorsal root ganglion repulsion
CC and growth cone collapse activities associated with semaphorins.
CC Polypeptides containing HTLS can be used to modulate the activity of
CC calcium channels. The peptides can also be used as an antigen to
CC generate antibodies that can then be used to modulate the activity of
CC calcium channels by inactivating naturally occurring channel ligands.
CC The peptides or antibodies can be used as drugs to treat any condition
CC or disease that is characterized by abnormal calcium channel function.

XX Sequence 771 AA;

Alignment Scores:

Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x AAG62726 (1-771)

QY	200	ATGGCTGGTAACTAGGATGCTGCTTTCTGGGAGTATTACTTACAGCAGACA	259
DB	1	MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla	20
QY	260	AACTATCAGAAATGGAAGCAATGTCCAGCGTGAATATCTCAAGAAATGTTG	319
DB	21	AsnTyrGlnAsnGlyAsnAsnValProArgLeuLeuSerTyrLysGluMetLeu	40
QY	320	GAATCCAAATGTGATCACTTCAATGCTTGGCCACAGCTCCAGTTATCATCCTTC	379
DB	41	GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe	60
QY	380	CTTTTGGATGAGAACGGATAGGCTGTATGTTGGAGCAAGGATCATATTTTCATTC	439
DB	61	LeuLeuAspGluGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe	80
QY	440	GACCTGGTTAATATCAAGATTTTCAAAGATTGTGCGCCAGTATCTTACACAGAGA	499
DB	81	AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg	100
QY	500	GATGAATCAAGTGGCTGGAAAGACATCTCTGAAAGAAATGTCTAAATTTTCATCAAGTA	559
DB	101	AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal	120
QY	560	CTTAAGGCATATATCAGACTCATTGTACGCTGTGGAACGGGGCTTTTCATCCCAATT	619
DB	121	LeuLysAlaLysAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle	140
QY	620	TGCACCTACATTGAAATTTGGACATCATCTCAGAGCAATATTTTAAAGCTGGAGAACTCA	679
DB	141	CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer	160
QY	680	CATTTTGAACCGCGCTGGGAAGAGTCCATATGACCTTAAGCTGCTGACAGCATCCCTT	739
DB	161	HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu	180
QY	740	TTATAGATGAGATTTACTCTGGAATCTGCACTGATTTTATGGGCGAGACTTTGCT	799
DB	181	LeuIleAspGlyLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla	200
QY	800	ATCTTCCAACTCTGGGCACACACCAATCAGACAGAGCAGCATGATTCCAGGTGG	859
DB	201	IlePheArgThrLeuGlyHisHisHisProIleArgThrGluGlnHisAspSerArgTrp	220
QY	860	CTCAATGATCAAAAGTTCAATTAGTGCCCACTCATCTCAGAGAGTGACAATCTCTGAAGAT	919

DB	221	LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp	240
QY	920	GACAAAGTATATCTTTTCTCCGTGAAAATCAATAGATGAGACACATCTCGAAAAGCT	979
DB	241	AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla	260
QY	980	ACTCAGCTAGAAATAGGTACATATGCAAGAAATGACTTTGGAGGCGACAGAAAGTCTGGTG	1039
DB	261	ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal	280
QY	1040	AATAAATGGACAACATCTCTCAAGCTGCTGATTTGCTCAGTCCAGGTCCTCAATGGC	1099
DB	281	AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly	300
QY	1100	ATTGACACTCATTTTGTGATGAACTGCAGATGATCTCTTAATGAATTTTAAAGATCCTAAA	1159
DB	301	IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys	320
QY	1160	AATCCAGTTGTATATGGAGTGTTTTACGACTTCCAGTAACATTTTCAAGGGATCAGCCGTG	1219
DB	321	AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal	340
QY	1220	TGATGTATAGCATGATGATGTGAGAGGTGTTCTTGGTCCATATGCCCACAGGAT	1279
DB	341	CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp	360
QY	1280	GGACCCAACTATCAATGGTGCCTTATCAAGGAGAGTCCCTATCCACGCCAGGAACT	1339
DB	361	GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr	380
QY	1340	TGTCCCAAGCAAAACATTTGGTGGTGTGACTCTCAAGAGACCTTCTCTGATGATGTATA	1399
DB	381	CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle	400
QY	1400	ACCTTTCCAAAGAGTATCCAGCCATCTACAATCCAGTGTTCCTATGAACAATCGCCCA	1459
DB	401	ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro	420
QY	1460	ATAGTGATCAAAACGGATGTAAATTTATCAATTTTACACAAATTTGCTGAGCCGAGTGGAT	1519
DB	421	IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValAlaAspArgValAsp	440
QY	1520	GCAGAAATGCAGATGATGATGTTTATGTTTATCGAAACAGATGTTGGACCGTCTTTAAA	1579
DB	441	AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys	460
QY	1580	GTAGTTTCAATTCCTAAGGAGACTTGGTATGATTTAGAAAGAGTTCCTGCGAAGAATG	1639
DB	461	ValValSerIleProLysGluThrTrpTyrAspLeuGluGluValLeuLeuGluGluMet	480
QY	1640	ACAGTTTTTTCGGNACCGACTGCTATTTTCAGCAATGGAGCTTCCACTAAGCAGCAACA	1699
DB	481	ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln	500
QY	1700	CTATATATTTGTTCAACGGCTGGGTTGCCAGCTCCCTTTTACACCGGTGTGATTTTAC	1759
DB	501	LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr	520
QY	1760	GGAAAGCGTGTGCTGATGTTGCTCTGCCCGAGACCTTTACTGCTGCTGGATGTTCT	1819
DB	521	GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaIatPaspGlySer	540
QY	1820	GCATGTTCTCGCTATTTTCCACTGCAAGCAGCAGCAGCAGCAAGATATAGAAT	1879
DB	541	AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn	560
QY	1880	GGAGACCCACTGACTCTTCTCAGACTTACACCATGATTAATCACCATGCGCCACAGCCCT	1939
DB	561	GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro	580
QY	1940	GAAGAGAGAAATCATCTATGCTGTAGAGAAATAGTAGCACAATTTTGGAAATCAGTCCGAAG	1999
DB	581	GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys	600

QY	2000	TCGCAGAGAGCGCTGCTCTATTGGCAATTTCCAGAGCGCGAAATAGAGAGCGAAAGAGAG	2059
Db	601	SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgAsnGluGluArgLysGluGlu	620
QY	2060	ATCAGAGTGGATGATCATATCATCATCAGCACAGATCAAGAGCGCTTCTGCTACGTAGTCTACAA	2119
Db	621	IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln	640
QY	2120	CAGAAGATTTCAGGCAATTAACCTCTGCCATCGCGTGGAAACATGGGTTTCATACAAACTCTT	2179
Db	641	GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu	660
QY	2180	CTTAAAGTAAACCTCGGAAGTCATTGACACAGAGCAATTTGGAAGCACTTCTCATAAAGAT	2239
Db	661	LeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLysAsp	680
QY	2240	GATGATGGAGATGCTCTTAGACCAAGAAATGTCCAATAGCATGACACCTAGCCAGAAG	2299
Db	681	AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys	700
QY	2300	GTCTGTGTACAGAGACTTCATGTCAGCTCATCAACCAACCCCAATCTCAACAGCATGGATGAG	2359
Db	701	ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu	720
QY	2360	TTCTGTGAACAAGTTTGGAAAGGACCGGAAACACACGTGGCAAGCGCCAGGACATACC	2419
Db	721	PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr	740
QY	2420	CCAGGGACAGTAAACAATGGAGCACTTACAGAAAAATAGAAAGTAGAACAACAGGAGG	2479
Db	741	ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg	760
QY	2480	ACCACCAATTTGAGAGCGCACCCAGGAGTGTC	2512
Db	761	ThrHisGluPheGluArgAlaProArgSerVal	771
RESULT 3			
ABG96413			
ID	ABG96413 standard; Protein; 771 AA.		
XX	AC ABG96413;		
XX	11-DEC-2002 (first entry)		
XX	Human ovarian cancer marker M473.		
XX	Human; ovarian cancer; marker; cancer; familial history; brain disorder;		
KW	central nervous system disorder; bacterial meningitis; viral meningitis;		
KW	Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;		
KW	brain herniation; inflammation; encephalitis; testicular disorder;		
KW	nontuberculous granulomatous orchitis; connective tissue disorder;		
KW	heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;		
XX	histological type; carcinogenic; ovarian cancer marker.		
XX	Homo sapiens.		
XX	W0200271928-A2.		
XX	19-SEP-2002.		
XX	14-MAR-2002; 2002WO-US07826.		
XX	14-MAR-2001; 2001US-276025P.		
PR	14-MAR-2001; 2001US-276026P.		
PR	10-AUG-2001; 2001US-311732P.		
PR	19-SEP-2001; 2001US-323580P.		
PR	26-SEP-2001; 2001US-324967P.		
PR	26-SEP-2001; 2001US-325102P.		
PR	26-SEP-2001; 2001US-325149P.		
XX	(MILL-) MILLENNIUM PHARM INC.		
XX	PA		
XX	XX		

QY 500 GATGAATGCAAGTGGCTGGAAAAGACATCTCTCAAGAAATGCTAAATTCATCAAGGTA 559
DB 101 AspGluCysLeuTrpAlaGlyAspIleLeuLysGluCysAlaAsnPhelIleLysVal 120
QY 560 CTTAAGGCATATAATCAGACTCACTTGATAGCTGTGGAAACGGGGCTTTTCATCAATTT 619
DB 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTAATGGAATCATCTCTGAGGCAATATTTTAAAGCTGAGAACTCA 679
DB 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAAGCGCGTGGGAAGAGTCCATATGACCCCTAAAGCTGTGACAGCATCCCTT 739
DB 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGGAATATTAATCTCTGGAAGTCTGAGTGTATTTATGGGGGAGACTTTGCT 799
DB 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCCGAACCTCTGGGACCAACCAATCAGACAGACAGCATGATTCAGGTGG 859
DB 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
QY 860 CTCATATGATCCAAAGTTCAATAGTGCACCTCATCTCAGAGATGCAAAATCCTGAAGAT 919
DB 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATATCTTTCTTCGCGTGAATAATGCAATAGATGGAGAACACTCTGGAAGACT 979
DB 241 AspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTAGATAGTACAGATGTCAGAAATGCTCTGATTTGCTCAGTCCAGGTCCAAATGGC 1039
DB 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGACAACTTCCTCAAGCTCGTCTGATTTGCTCAGTCCAGGTCCAAATGGC 1099
DB 281 AsnLysTrpThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTATTTGATGAATGCTGAGGATGATTTCTTAATGAATCTTAAAGATCCTAAA 1159
DB 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCAGTTGTATATGAGTGTTCAGCTTCCAGTCAATTTTCAAGGGATCAGCCGTG 1219
DB 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGATGTGAGAAGGTGTTCTTGTGTCATATGCCACAGGGAT 1279
DB 341 CysMetTyrSerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCAACTATCAATGGGTGCTTATCAAGGAAGAGTCCCTATTCACGGCCAGGAAT 1339
DB 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCACGACAAACATTTGGTGTTCACCTCAAAAGGACCTTCTCTGATGATGTTATA 1399
DB 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspValIle 400
QY 1400 ACCTTTGCAAGAAGTCATCCAGCATGTACATCCAGTGTTCCTTATGAAACAATCGCCCA 1459
DB 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTGATCAAAACGGATTAATTAATCAATTTACAAATTTGCTGATAGCCGAGTGGAT 1519
DB 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAGATGACAGTATGATGTTATGTTATCGGAACAGATGTTGGGACCGTCTTAAA 1579
DB 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTTCTTAAGGAGACTTGGTATGATTTAGAAGAGGTTCCTGCTGGAAGAAATG 1639

DB 461 ValValSerIleProLysGluThrTyrAspLeuGluValLeuLeuGluGluMet 480
QY 1640 ACAGTTTTTCCGGAACCGACTGCTATTTCACCAATGAGAGCTTTCACCAAGCAGCAACA 1699
DB 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
QY 1700 CTATATATGTTTCAACGGCTGGGGTTCGCCAGCTCCCTTTACACCGGTGTGATATTTAC 1759
DB 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGCGTGTGCTGAGTGTTCCTGCCCGAGACCCCTTACTGCTGTGGATGGTTCT 1819
DB 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCACTGCAAGAGACGACAGACCAAGATATTAAGAAAT 1879
DB 541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgArgGlnAspIleArgAsn 560
QY 1880 GGAGACCACTGACTCACTGTTTCAGACTTACACCATGATATCACCATGGCCACAGCCCT 1939
DB 561 GlyAspProLeuThrHisCysSerAspPheHisAspAsnHisHisGlyHisSerPro 580
QY 1940 GAAGAGAGAATCATCTATGTTGTAGAGAATAGTAGACATTTTGGAAATGCAGTCCGAAG 1999
DB 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGAGCGTGTGCTTATTTGGCAATTCAGAGGGGCAATTAAGAGCGCAAGAGAG 2059
DB 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgAsnGluGluArgLysGluGlu 620
QY 2060 ATCAGATGATGATCATATCATCAGACAGATCAAGCGCTTCTGCTACGTAGTCTACAA 2119
DB 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAAGGATTCAGCAATTTACCTCTGCCATCGCGTGGAAACATGGGTTTCATACAACTCTT 2179
DB 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGTAACTCCGGAAGTCATTGACACAGAGCATTTGGNAGAACTTCTTCATAAAGAT 2239
DB 661 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLysAsp 680
QY 2240 GATCATCGAGATGCTCTTAAGACCAAGAAATGTCCAAATAGCATGACACCTAGCCAGAG 2299
DB 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGTACAGACTTTCATCTCAGCTCATCAACCCACCCCAATCTCAACACGATGATGAG 2359
DB 701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
QY 2360 TTCTGTGAACAGTTTGGAAAAGGGACCGAAACAACTCGGCAAGCGCCAGGACATACC 2419
DB 721 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAGGGAACAGTAACAATGGAAGCACTTCAAGAAAATTAAGAAAGGTAGAAAACAGAGG 2479
DB 741 ProGlyAsnSerAsnLysTrpLysHisLeuGlnLysLysLysGlyArgAsnArgArg 760
QY 2480 ACCCAGCAATTTGAGAGGGCACCCAGGAGTGTCT 2512
DB 761 ThrHisGluPheGluArgAlaProArgSerVal 771
RESULT 4
AAY21264
ID AAY21264 standard; Protein; 796 AA.
XX AC AAY21264;
XX AC AAY21264;
XX DT 22-JUL-1999 (first entry)
XX DE Human semaphorin III wild type protein fragment 1.
XX

KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Homo sapiens.

XX
 XX WO9845322-A2.
 XX
 XX PD 15-OCT-1998.
 XX
 XX PF 02-APR-1998; 98WO-IB00705.
 XX
 XX PR 10-APR-1997; 97US-0043163.
 XX
 XX PA (UYUT-) RIJKSUNIV UTRECHT.
 XX
 XX PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 XX
 XX PA (UYRO-) UNIV ROTTERDAM ERASMUS.

XX
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
 XX
 XX WPI; 1998-609901/51.
 DR
 DR N-PSDB; AAX75767.

XX
 XX Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 XX Disclosure; Figure 16; 258pp; English.

PS This invention describes a novel method for the diagnosis of a disease
 XX caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.
 XX
 XX Sequence 796 AA;

Alignment Scores:
 Pred. No.: 0 Length: 796
 Score: 437.00 Matches: 437
 Percent Similarity: 100.00% Conservativity: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 48.83% Indels: 0
 DB: 19 Gaps: 0

US-09-774-490-1 (1-2709) x AAY21264 (1-796)

OY 194 TGCAGCTGGCTGGTAACTAGGATTCGTCTTTCTGGGGAGTATTACTACGCA 253
 DB 4 CysSerMetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAla 23

OY 254 AGAGCAAACTATCAGAAATGGGAAGAACAAATGTGCCAAGGCTGAAATTTATCTACAAAGAA 313
 DB 24 ArgAlaAsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGlu 43
 OY 314 ATGTTGGAAATCCAAACATGTGATCACTTTCATGGCTTGGCCACAGCTCCAGTTATCAT 373
 DB 44 MetLeuGluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHis 63
 OY 374 ACCTTCTCTTTGGATGAGAAACGAGTACGCTGTATGTTGGAGCAAAAGGATCACATATTT 433
 DB 64 ThrPheLeuLeuAspGluArgSerArgLeuTyrValGlyAlaLysAspHisIlePhe 83
 OY 434 TCATTGCACTGGTTAATATCAAGATTTTCAAAAGATTGTGTGCCAGTATCTTACACC 493
 DB 84 SerPheAspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThr 103
 OY 494 AGAAGAGATGAATCAAGTGGCTGGAAAGACACATCTCTGAAGATGTGCTAAATTTTCATC 553
 DB 104 ArgArgAspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIle 123
 OY 554 AAGGTACTTAAAGGCATATATCAGACTCACCTTGTACGCTGTGGAAACGGGGCTTTTCAT 613
 DB 124 LysValLeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHis 143
 OY 614 CCAATTTGCACCTACATTGGAATTTGGACATCATCTCTGAGGACAATATTTTAAAGTCGAG 673
 DB 144 ProIleCysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGlu 163
 OY 674 AACTCACATTTTGAACGGCTGGGAGAGTCCATATGACCCCTAAGCTCTCTGACACA 733
 DB 164 AsnSerHisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAla 183
 OY 734 TCCCTTTTAAATAGATGAGAAATTTACTCTCGAACTGCAGCTGATTTTATGGGGAGAC 793
 DB 184 SerLeuLeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAsp 203
 OY 794 TTTGCTATCTTCCGAACCTTTTGGGCACACACCAATCAGGACAGACAGCATGATTC 853
 DB 204 PheAlaIlePheArgThrLeuGlyHisHisHisProIleArgThrGluGlnHisAspSer 223
 OY 854 AGTGGCTCAATGATCCAAAGTTTCATTAGTGGCCACCTCATCTCAGAGAGTGCATCTCT 913
 DB 224 ArgTrpLeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnPro 243
 OY 914 GAAGATCAAAAGTATATCTTTCTTCGTGAAATGCAATAGATGGAGAACACTCTCGA 973
 DB 244 GluAspAspLysValTyrPhePheArgGluAsnAlaIleAspGlyGluHisSerGly 263
 OY 974 AAAGCTACTCAGCTAGATAGGTGAGATATGCAAGATGACTTTTGGAGGGCAGAGAGT 1033
 DB 264 LysAlaThrHisAlaAArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSer 283
 OY 1034 CTGTTGTAATAATGGACAACATTTCTCAAAGCTGCTGATTTGCTCAGTCCAGGTCCA 1093
 DB 284 LeuValAsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyPro 303
 OY 1094 AATGGCATTCACACTCATTTTGTGAATGCTGAGATGTATTCTTAATGAACCTTTAAAGAT 1153
 DB 304 AsnGlyIleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPhenylsAsp 323
 OY 1154 CCTAAAAATCCAGTTGTATAGGAGTGTTCACGACTTCCAGTAAACATTTTCAAGGGATCA 1213
 DB 324 ProLysAsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySer 343
 OY 1214 GCCGTGTATGTATACGATGATGTGAGAAGGTGTTCCTTGGTCCATATGCCAC 1273
 DB 344 AlaValCysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHis 363
 OY 1274 AGGATGACCCCACTATCAATGGTGCCTTTATCAAGGAAGAGTCCCTCTTCCACGGCCA 1333
 DB 364 ArgAspGlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgPro 383
 OY 1334 GGAACCTTGTCCCGCAAAACATTTTGGTGGTGTGACTCTCAAGAGGACCTTCTCTGATGAT 1393

Db 384 GlyThrCysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAsp 403
 Qy 1394 GTTATAACCTTTGCAAGAAGTCATCCAGCATGTACATCCAGTGTTCCTATGAAACAAT 1453
 Db 404 ValIleThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsn 423
 Qy 1454 CCGCCAATAGTCATCAAAACGATGTAAATTATCAATTATACACAAATTGTC 1504
 Db 424 ArgProIleValIleLysThrAspValAsnTyrGlnPheThrGlnIleVal 440

RESULT 5

AAR74175

ID AAR74175 standard; Protein; 477 AA.

XX

AC AAR74175;

XX

DT 01-NOV-1995 (first entry)

XX

DE Human collapsin.

XX

KW Collapsin; antibody; therapy.

XX

OS Homo sapiens.

XX

FH Key Location/Qualifiers

FT Binding-site 9..19

FT Binding-site /note= "antibody binding site"

FT Binding-site 51..65

FT Binding-site /note= "antibody binding site"

XX

PN US5416197-A.

XX

PD 16-MAY-1995.

XX

PF 15-OCT-1993; 93US-0136922.

XX

PR 15-OCT-1993; 93US-0136922.

XX

PA (UYPE-) UNIV PENNSYLVANIA.

XX

PI Luo Y, Raper JA;

XX

XX WPI; 1995-193478/25.

DR N-PSDB; A9Q92331.

XX

PT New antibody to human collapsin - used to inhibit the activity of

PT collapsin, to induce neurite out-growth and to treat individuals with

PT nerve damage.

XX

XX Claim 2; Columns 15-18; 11pp; English.

XX

CC An antibody capable of specifically binding at least a portion of

CC the collapsin protein can be used to purify human collapsin and

CC to inhibit the activity of the protein. It can be used to induce

CC neurite outgrowth by neuronal cells and to treat individuals

CC suffering from nerve damage.

XX

SQ Sequence 477 AA;

XX

XX Alignment Scores:

XX

XX Pred. No.:

XX Score:

XX Length:

XX 434.00

XX Matches:

XX 434

XX Percent Similarity:

XX 100.00%

XX Best Local Similarity:

XX 100.00%

XX Query Match:

XX 48.49%

XX Indels:

XX 0

XX Gaps:

XX 16

XX DB:

XX 16

XX US-09-774-490-1 (1-2709) x AAR74175 (1-477)

XX

Qy 905 GACATCTCGAGATGACAAATGATACCTTTCTTCCTGAAATGCAATGATGGAGAA 964

Db 4 AspAsnProGluAspAspLysValTyrPhePheArgGluAsnAlaIleAspGlyGlu 23

Qy 965 CACTCTGAAAGCTACTCACGCTAGATAGTGCAGATATGCAAGAATGACTTTGAGGG 1024
 Db 24 HisSerGlyLysAlaThrHisAlaArgLleGlyGlnIleCysLysAsnAspPheGlyGly 43
 Qy 1025 CACAGAAGTCTGCTGTAATAAATGACACAACTTCCTCAAGCTCGTCTGATTGCTCAGTG 1084
 Db 44 HisArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeuLleCysSerVal 63
 Qy 1085 CCAGGTCAAATGGCATTGACACTCATTTTGTATGATGAACCTGCAGGATGATTTCTTAATGAAC 1144
 Db 64 ProGlyProAsnGlyIleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsn 83
 Qy 1145 TTTAAAGATCCTAAATATCCAGTTGTATATGAGTGTTCAGTCTTCAGTCTTCAGTAACTTTTC 1204
 Db 84 PheLysAspProLysAsnProValValTyrGlyValPheThrThrSerSerAsnLlePhe 103
 Qy 1205 AAGGATCAGCCGTGTGTATGTATAGTATGATGATGTGAGAAGGGTGTTCCTTGGTCCA 1264
 Db 104 LysGlySerAlaValCysMetTyrSerMetSerAspValArgValPheLeuGlyPro 123
 Qy 1265 TATCCCAACAGGATGACACCACTATCAATGGTGTCTTATCAAGGAAGAGTCCCTAT 1324
 Db 124 TyrAlaHisArgAspGlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyr 143
 Qy 1325 CCACGGCCAGGAACCTTCTCCAGCAAAACATTTTGGTGTGTTCAGTCTTCACAAAGACCTT 1384
 Db 144 ProArgProGlyThrCysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeu 163
 Qy 1385 CCTGATGATGTATTAACCTTTGCAAGAAAGTCAATCCAGCCATGTACAATCCAGTGTTCCT 1444
 Db 164 ProAspAspValIleThrPheAlaArgSerHisProAlaMetTyrAsnProValPhePro 183
 Qy 1445 ATGAACAATGCCCAATAGTATCAAAACCGATGTAATTAATTAATTAATTAATTAATTAAT 1504
 Db 184 MetAsnAsnArgProIleValIleLysThrAspValAsnTyrGlnPheThrGlnIleVal 203
 Qy 1505 GTAGACCGAGTGGATGACAGAAGATGACAGTATGATGTATGTTTATCGGAACAGATGTT 1564
 Db 204 ValAspArgValAspAlaGluAspGlyGlnTyrAspValMetPheLleGlyThrAspVal 223
 Qy 1565 GGGACCGTCTTAAAGTAGTTTCAATTCCTAAGAGAGACTTGGTATGATTTAGAGAGGTT 1624
 Db 224 GlyThrValLeuLysValValSerIleProLysGluThrTrpTyrAspLeuGluVal 243
 Qy 1625 CTGCTGAAGAATGACAGTCTTTCGGGAACCGACTCTCTATTTTCAGCAATGGAGCTTCC 1684
 Db 244 LeuLeuGluGluMetThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSer 263
 Qy 1685 ACTAAGCAGCAACAACCTATATATTGTTTCAACGCTGGGTTGCCAGCTCCCTTTTACAC 1744
 Db 264 ThrLysGlnGlnLeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHis 283
 Qy 1745 CGGTGTGATATTTACGGGAAAGCGTGTGCTGAGTGTTCCTCGCCCGAGACCTTACTGT 1804
 Db 284 ArgCysAspIleTyrGlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrCys 303
 Qy 1805 GCTTGGATGTTCTGCATGTTCTCGCTATTTTCCACTGCAAGAGACCGCAGACGCA 1864
 Db 304 AlaTrpAspGlySerAlaCysSerArgTyrPheProThrAlaLysArgThrArgArg 323
 Qy 1865 CAAGATATAAGAAATGGAGACCCACTGACTCTGTTTTCAGCTTACACCATGATAATCAC 1924
 Db 324 GlnAspIleArgAsnGlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHis 343
 Qy 1925 CATGGCCACACCCCTGAAGAGAGAATCATCTATGGTGTAGAGATAGTAGACATTTTGT 1984
 Db 344 HisGlyHisSerProGluGluArgLleIleTyrGlyValGluAsnSerSerThrPheLeu 363
 Qy 1985 GAATGAGTCCGAGTCCGAGAGAGCGCTGCTCTATTTGGCAATTCAGAGCGGAAATGAA 2044
 Db 364 GluCysSerProLysSerGlnArgAlaLeuValTyrTrpGlnPheGlnArgAsnGlu 383

QY 2045 GAGCGAAAGACAGATCAGAGTGTGATCATATCATCATCAGACAGATCAAGGCTTCTG 2104
 Db |||||||||||||||||||
 384 GluArgLysGluGluIleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeu 403
 |||||||||||||||||||
 QY 2105 CTACGTAGTCTACACAGAGATTGAGGCAATTAACCTCTGCATGCGGTGGAACATCGG 2164
 Db |||||||||||||||||||
 404 LeuArgSerLeuGlnGlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGly 423
 |||||||||||||||||||
 QY 2165 TTATACAACTCTTCTTAAGTAAACCTCGGAGTCAATTCAC 2206
 Db |||||||||||||||||||
 424 PheIleGlnThrLeuLeuLysValThrLeuGluValIleAsp 437
 |||||||||||||||||||
 RESULT 6
 ID ABG54800 standard; Peptide; 151 AA.
 XX
 AC ABG54800;
 XX
 DT 25-FEB-2003 (first entry)
 DE
 DE Human liver peptide, SEQ ID No 33448.
 XX
 KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200157273-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00664.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488898/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analysing gene expression in human adult liver -
 XX
 PS Claim 27; SEQ ID No 33448; 658pp; English.
 XX
 CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult
 CC liver. (I) may be used for predicting, measuring and displaying gene
 CC expression in samples derived from human adult liver. The genes
 CC identified may be involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
 CC is associated with coronary heart disease. ABG47348-ABG59930 represent
 CC human liver single exon encoded peptides of the invention.
 CC Note: The sequence information for this patent does not appear in the
 CC printed specification but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 151 AA;
 Alignment Scores:
 Pred. No.: 7,28e-145 Length: 151
 Score: 151.00 Matches: 151
 Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 16.87% Indels: 0
 DB: 22 Gaps: 0
 US-09-774-490-1 (1-2709) x ABG54800 (1-151)
 QY 2060 ATCAGATGGATGATCATATCATCATCAGACAGATCAAGGCTTCTGCTAGTGTCTACAA 2119
 Db 1 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 20
 |||||||||||||||||||
 QY 2120 CAGAGAGATTGAGGCAATTAACCTCTGCATGCGGTGGAACATCGGTTTCTACAACTCTT 2179
 Db 21 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGluHisGlyPheIleGlnThrLeu 40
 |||||||||||||||||||
 QY 2180 CTTAAGGTAAACCTCGGAGTCAATTCACACAGAGATTTGGAGAAGAACTTCTTCATAAAGAT 2239
 Db 41 LeuLysValThrLeuGluValIleAspThrGluHisLeuGluLeuLeuHisLysAsp 60
 |||||||||||||||||||
 QY 2240 GATGATGAGATGCTCTTAAGACCAAGAAATGTCCAATACATGACACCTAGCCAGAAAG 2299
 Db 61 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 80
 |||||||||||||||||||
 QY 2300 GTCTGTACAGAGACTTCATGACGATCATCAACACCCCATCTCAACACGATGATCAG 2359
 Db 81 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 100
 |||||||||||||||||||
 QY 2360 TTCTGTCAACAAAGTTTGGAAAAGGACCGAAAACAAAGCTCGGCAAGCCAGGACATACC 2419
 Db 101 PheCysGluGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 120
 |||||||||||||||||||
 QY 2420 CCAGGGAACAGTAACAAATGGAAGCATTCAAGAAATAAGAAAGGTAGAAAACAGGAGG 2479
 Db 121 ProGlyAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 140
 |||||||||||||||||||
 QY 2480 ACCCAGCAATTTGAGGGGACCCAGGAGTCTC 2512
 Db 141 ThrHisGluPheGluArgAlaProArgSerVal 151
 |||||||||||||||||||
 RESULT 7
 ID ABG42930
 XX
 AC ABG42930;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 32595.
 KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Rudik syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00665.
 XX
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.

CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease, II
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.

XX Sequence 81 AA;

XX Alignment Scores:

Pred. No.:	2,81e-73	Length:	81
Score:	81.00	Matches:	81
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	9.05%	Indels:	0
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) x AAY21293 (1-81)

QY 2286 CACCTAGCCAGAGGCTGTGTACAGAGCTTCATGACCTCAACACCCCAATTCGA 2345
 Db 1 HistenuAaArgArgSerGlyThrGluThrSerCySseSerSerThrThroIleSer 20
 QY 2346 ACACGATGATGATGTTCTGTGAACAAGTTGAAAAAGGACCCGAAACAACTGCGCAA 2405
 Db 21 ThrArgTpmSeSerSerValasnlySpheGlyLyGlyThGluAsnAsnValGlyLyS 40
 QY 2406 GGGCCAGCATATCCCGAGGAGACATTAACAAATGACAGCATTACAGAAATAAGAAAG 2465
 Db 41 GlyGlnAspIleProGlnGlyThrValThrAsnGlySerThrTyrLysLysIleArgLyS 60
 QY 2466 GTAGAAACAGAGAGACCCAGAAATTTGAGAGGACCCAGAGATGTCAGCTGCAATTA 2525
 Db 61 ValGluThrGlyGlyProThrAsnLeuArgGlyHisProGlyValSerGluLeuHisTyr 80
 QY 2526 CTC 2528
 Db 81 Leu 81

RESULT 9

AAY21288 standard; Protein; 77 AA.

XX AAY21288;

XX 22-JUL-1999 (first entry)

XX Human semaphorin III mutant protein fragment 21.

XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 XX frameshift mutation; age-related disease; neurodegenerative disorder;
 XX Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 XX Huntington's disease; multiple sclerosis; alcoholic liver disease;
 XX diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 XX ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 XX neurofilament-F; presentin I; presentin II; cellular tumour antigen;
 XX glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 XX bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSF-A;
 XX high mobility group protein-C; neuroendocrine specific protein A.
 XX Synthetic.

OS Homo sapiens.
 XX WO9845322-A2.
 PN 15-OCT-1998.
 XX 02-APR-1998; 98WO-IB00705.
 PF 10-APR-1997; 97US-0043163.
 PR (UYT-) RIJCKINIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX Burdach JPH, Grosfeld FG, Van Leeuwen FW;
 DR WPT, 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 RNA
 Disclosure; Figure 16; 258pp; English.

CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.

XX Sequence 77 AA;

Alignment Scores:	3,47e-69	Length:	77
Pred. No.:	77.00	Matches:	77
Score:	77.00	Conservative:	0
Percent Similarity:	100.00%	Mismatches:	0
Best Local Similarity:	100.00%	Indels:	0
Query Match:	8.60%	Gaps:	0
DB:	19		

US-09-774-490-1 (1-2709) x AAY21288 (1-77)

QY 1641 CAGTTTTCGGGAAACCGACTCTATTTCAGCAATGAGCTTTCACATAGACGACAAAC 1700
 Db 1 GlnPhePheGlyAsnArgLeuLeuPheGlnGlnIlnTrpSerPheProLeuSerSerAsn 20
 QY 1701 TATATATTGTTTCAACGGCTGGGTGGCCAGCTCCCTTACACCGGCTGATATTACG 1760
 Db 21 TyrIleLeuValGlnArgLeuGlyLeuProSerLeuTyrTrnGlyValIlePheThr 40
 QY 1761 GGAAGCGTGTGAGTGGTGGCTCGCCGACGACCCCTTACTGTGCTGGAGTGGTCTG 1820
 Db 41 GlyLysArgValIleSerValAlaSerProGluThrLeuTrnValLeuGlyMetValLeu 60
 QY 1821 CATGTTCTCGCTATTTCCTCACTGCACAGACGACGACGACGACGACGACGATA 1871
 Db 61 HisValIleuAlaIlePheProLeuGlnArgGspAlaGlnAspAlaPlySile 77

RESULT 10
AAV21291
ID AAV21291 standard; Protein; 74 AA.

AC AAV21291;

DT 22-JUL-1999 (first entry)

DE Human semaphorin III mutant protein fragment 24.

XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
KW frameshift mutation; age-related disease; neurodegenerative disorder;
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
KW high mobility group protein-C; neuroendocrine specific protein A.

OS Synthetic.
OS Homo sapiens.

PN WO9845322-A2.

XX 15-OCT-1998.

XX 02-APR-1998; 98WO-IB00705.

XX 10-APR-1997; 97US-0043163.

XX (UYUT-) RIKSUNIV UTRECHT.

PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
(UYRO-) UNIV ROTTERDAM ERASMUS.

PI Burbach JPH, Grosveld FG, Van Leeuwen FW;

DR MPI; 1998-609901/51.

DR N-PSDB; AAX75767.

PT Diagnosing disease by detecting frameshift mutations in RNA or
PT corresponding protein mutations - used to diagnose cancer and
PT neurological diseases, particularly Alzheimer's disease, and also
PT for treatment and prevention with specific ribozymes or wild-type
PT RNA

PS Disclosure; Figure 16; 258pp; English.

XX This invention describes a novel method for the diagnosis of a disease
XX caused by, or associated with, an RNA molecule that has a frameshift
XX mutation. The method is used to diagnose age-related diseases, especially
XX cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
XX disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
XX multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
XX and many others listed) or susceptibility to these disorders. The method
XX allows a definitive diagnosis of Alzheimer's disease in living patients,
XX at an early stage. It is based on the observation that disease may be
XX caused by mutations in RNA rather than DNA. The invention describes the
XX use of neuronal system RNA molecules, specifically proteins including
XX beta-amyloid precursor protein (beta-APP), the microtubule associated
XX proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
XX associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
XX neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
XX protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
XX 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
XX protein-C (HMGP-C) and neuroendocrine specific protein A.

XX Sequence 74 AA;

Alignment Scores: 4.07e-66 Length: 74
Pred. No.:

Score: 74.00 Matches: 74
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.27% Indels: 0
DB: 19 Gaps: 0

US-09-774-490-1 (1-2709) x AAV21291 (1-74)

QY 1965 AGAATAGTAGCATTGTTGGATGACGCGAGTGGCAGCGCTGCTATTGGC 2024

DB 1 ArgilleValAlaHisPheTrpAsnAlaValAlaGserArgAlaGlnArgTrpSerIleGly 20

QY 2025 AATTCAGAGCGCAATGAAGCGCAAGAAAGAGATCAGATGATCATATCATCA 2084

DB 21 AanserArgGlyGluMetIysSerGluIysArgSerGluTrpMetIleIleSer 40

QY 2085 GCACAGATCAGAGCCTTGTCTAGCTAGTCTTCAACAGAGGATTCAGCAATTAACCTCT 2144

DB 41 GlyGlnIleLysAlaPheCysTyrValValTyrAsnArgArgIleGlnAlaIleThrSer 60

QY 2145 GCCATGCGGTGGACATGGCTCATACAAACCTCTTAAGG 2186

DB 61 AlaMetArgTrpAsnMetGlySerTyrIysLeuPheLeuArg 74

RESULT 11

AAV21274
ID AAV21274 standard; Protein; 69 AA.

XX AAV21274;

XX 22-JUL-1999 (first entry)

DE Human semaphorin III mutant protein fragment 7.

XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
KW frameshift mutation; age-related disease; neurodegenerative disorder;
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
KW high mobility group protein-C; neuroendocrine specific protein A.

OS Synthetic.

OS Homo sapiens.

PN WO9845322-A2.

XX 15-OCT-1998.

XX 02-APR-1998; 98WO-IB00705.

XX 10-APR-1997; 97US-0043163.

PA (UYUT-) RIKSUNIV UTRECHT.

PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
(UYRO-) UNIV ROTTERDAM ERASMUS.

PI Burbach JPH, Grosveld FG, Van Leeuwen FW;

DR MPI; 1998-609901/51.

DR N-PSDB; AAX75767.

PT Diagnosing disease by detecting frameshift mutations in RNA or
PT corresponding protein mutations - used to diagnose cancer and
PT neurological diseases, particularly Alzheimer's disease, and also
PT for treatment and prevention with specific ribozymes or wild-type
PT RNA

PS Disclosure; Figure 16; 258pp; English.

CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease, II
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-Ap), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HRP-1, high mobility group
 CC protein-C (HMGp-C) and neuroendocrine specific protein A.

XX SQ Sequence 69 AA;

Alignment Scores:

Pred. No.: 5,3e-61 Length: 69
 Score: 69.00 Matches: 69
 Percent Similarity: 100.00% Conservatve: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 7.71% Indels: 0
 DB: 19 Gaps: 0

US-09-774-490-1 (1-2709) x AAY21274 (1-69)

QY 747 ATGAGAAATTAATCTCTGGAAGTGAAGTATTTATGGGGGAGACTTGTATCTTCC 806
 DB 1 MetGluuAnTyThrLeuGluuGluuGluuLeuTrpGluThrLeuLeuSerSer 20
 QY 807 GAACCTTTGGGACCAACCAACCAATGAGACAGACAGATATTCAGTGGCTCAATG 866
 DB 21 GluLeuLeuGlyThrThrThrGlnSerGlyGlnSerSerMetLeProGlyGlySerMet 40
 QY 867 ATCCAAAGTCAATGATGTCGCCACCTCATCTCAGACAGACAGCAATCTGGAAGTGAAGAAG 926
 DB 41 IlegInSerSerLeuValProThrSerSerGlnArgValThrLeuLeuMetThrIlys 60
 QY 927 TATACCTTTCTTCGCGTGAATGCAA 953
 DB 61 TyrThrPheSerSerValIysMetGln 69

RESULT 12

AAY21272 standard; Protein; 64 AA.

XX AAY21272;

DT 22-UTU-1999 (first entry)

DE Human semaphorin III mutant protein fragment 5.

XX Human, beta-amyloid precursor protein; beta-Ap; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presentin I; presentin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HRP-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGp-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.

OS Synthetic.
 OS Homo sapiens.

XX WO9845322-A2.

XX

PD 15-OCT-1998.

XX 02-APR-1998; 98WO-IB00705.

XX 10-APR-1997; 97US-0043163.

XX (UYUT-) RIJCKSUNIV UTRECHT.

XX (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.

XX (UYRO-) UNIV ROTTERDAM ERASMUS.

XX Burbach JPH, Grosveld FG, Van Leeuwen FW;

XX WPI; 1998-609901/51.

XX N-PSDB; AAX75767.

PT Diagnosing disease by detecting frameshift mutations in RNA or

PT corresponding protein mutations - used to diagnose cancer and

PT neurological diseases, particularly Alzheimer's disease, and also

PT for treatment and prevention with specific ribozymes or wild-type

PT RNA

PS Disclosure; Figure 16; 258pp; English.

CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease, II
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-Ap), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HRP-1, high mobility group
 CC protein-C (HMGp-C) and neuroendocrine specific protein A.

XX SQ Sequence 64 AA;

Alignment Scores:

Pred. No.: 6,91e-56 Length: 64
 Score: 64.00 Matches: 64
 Percent Similarity: 100.00% Conservatve: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 7.15% Indels: 0
 DB: 19 Gaps: 0

US-09-774-490-1 (1-2709) x AAY21272 (1-64)

QY 534 AAGATGTGCTAATTTATCAAGGATTAAGCATATATACAGACTTCATGTCGCT 593

DB 1 LysAsnValLeuIleSerSerArgIleuArgHisIleIleArgLeuThrCysThrPro 20

QY 594 GTGAAAGCGGCGCTTTTCATCCATTTGACCTACATTTGAATTTGACATCATCTGAGG 653

DB 21 ValGluArgGlyLeuPheIleGlnPheAlaProThrLeuIleuValAspIleIleuArg 40

QY 654 ACAATATTTTAAAGTGAAGAACTCACTTTTGAAGAAAGCGCGGGAAGAGTCCATATG 713

DB 41 ThrIlePheLeuSerThrIleArgThrHisIleLeuIleuValAlaValArgValHisMet 60

QY 714 ACCCTAAGCTGC 725

DB 61 ThrLeuSerCys 64

RESULT 13

AAY21299 standard; Protein; 57 AA.

XX

XX AAY21299;
 AC
 XX
 DT 22-JUL-1999 (first entry)
 DE Human semaphorin III mutant protein fragment 32.
 XX
 XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KM frameshift mutation; age-related disease; neurodegenerative disorder;
 KM Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KM Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KM diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KM ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KM neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KM glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KM bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KM high mobility group protein-C; neuroendocrine specific protein A.
 XX
 XX Synthetic.
 OS Homo sapiens.
 OS
 PN MO9845322-A2.
 PD 15-OCT-1998.
 PF 02-APR-1998; 98WO-IB00705.
 PR 10-APR-1997; 97US-0043163.
 XX
 XX (UYUT-) RIKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
 XX WPI; 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 PS Disclosure; Figure 16; 258pp; English.
 XX
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.
 XX
 XX Sequence 57 AA;
 SQ
 Alignment Scores:
 Pred. No.: 16-48 Length: 57
 Score: 57.00 Matches: 57
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 6.37% Indels: 0

DB: 19 Gaps: 0
 US-09-774-490-1 (1-2709) x AAY21299 (1-57)
 QY 217 GATGTCGTCTTTCTGGGAGTATTACTTACACGACGAACTACAGATGGAA 276
 DB 1 AspCysIeuSerPheIeuNylYserIleThrYrSerIlySerIlySerIlySerIlyTrpGlu 20
 QY 277 GAACATATGCCAAGGCTGAATTTATCCATCAAGAAAGTGGAAATCAAGATGAT 336
 DB 21 GluGlnCysAlaIysAlaIleIleIleuGlnIleArgAsnValGlyIleGlnIlyCysAsp 40
 QY 337 CACTTCAATGGCTTGGCCACAGCTCCAGTTATCATCTTCTTTTGA 387
 DB 41 HisPheGlnTrpIleuNylGlnIleuGlnIleuSerIlyLeuProPheIly 57
 RESULT 14
 AAY21271
 ID AAY21271 standard; Protein; 85 AA.
 XX
 AC AAY21271;
 XX
 DT 22-JUL-1999 (first entry)
 DE Human semaphorin III mutant protein fragment 4.
 XX
 XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KM frameshift mutation; age-related disease; neurodegenerative disorder;
 KM Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KM Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KM diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KM ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KM neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KM glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KM bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KM high mobility group protein-C; neuroendocrine specific protein A.
 XX
 XX Synthetic.
 OS Homo sapiens.
 OS
 PN MO9845322-A2.
 PD 15-OCT-1998.
 PF 02-APR-1998; 98WO-IB00705.
 PR 10-APR-1997; 97US-0043163.
 XX
 XX (UYUT-) RIKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
 XX WPI; 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 PS Disclosure; Figure 16; 258pp; English.
 XX
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,

CC at an early stage. It is based on the observation that disease may be
CC caused by mutations in RNA rather than DNA. The invention describes the
CC use of neuronal system RNA molecules, specifically proteins including the
CC beta-amyloid precursor protein (beta-APP), the microtubule associated
CC protein Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
CC neurofilament-F, presenilin 1, presenilin II, glial fibrillary acidic
CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
CC protein-C (HMGp-C) and neuroendocrine specific protein A.

XX Sequence 85 AA;

Alignment Scores:

Pred. No.:	1,05e-46	Length:	85
Score:	55.00	Matches:	55
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	6.1%	Indels:	0
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) x AAY21271 (1-85)

QY 366 GTTATCATACCTTCTTTGGATGAGGAAAGGAGTGGTGTGAGCAAGATC 425
|||
Db 31 ValIleIleProSerPheTrpMetArgAsnGlyValGlyCysMetLeuGluGlnArgIle 50

QY 426 ACATATTTTCATTCGACCTGGTTAATATCAAGATTTTCAAAAGATGTGCGCCATAT 485
|||
Db 51 ThrTrpPheHisSerThrTrpLeuIleSerArgIlePheLeuYargLeuGlyGlnIlyr 70

QY 486 CTTACACCAAGAGATGATGATGCAAGTGGGCTGCAAAAGATCC 530
|||
Db 71 LeuThrProGluGluMetAsnAlaSerGlyLeuGluIlyThrSer 85

RESULT 15

ABG54764
ID ABG54764 standard; Peptide; 57 AA.

XX AC ABG54764;

DT 25-FEB-2003 (first entry)

XX Human liver peptide, SEQ ID No 33412.

DE Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;

KW hypercholesterolaemia; coronary heart disease.

XX Homo sapiens.

XX WO200157273-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00664.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488898/53.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX PT analysing gene expression in human adult liver -

PS Claim 27; SEQ ID No 33412; 658bp; English.

XX The invention relates to a single exon nucleic acid probe (SRNP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult
CC liver. (I) may be used for predicting, measuring and displaying gene
CC expression in samples derived from human adult liver. The genes
CC identified may be involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
CC is associated with coronary heart disease. ABG47348-ABG5930 represent
CC human liver single exon encoded peptides of the invention.
CC Note: The sequence information for this patent does not appear in the
CC printed specification but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 57 AA;

Alignment Scores:			
Pred. No.:	1.23e-44	Length:	57
Score:	53.00	Matches:	53
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.92%	Indels:	0
DB:	22	Gaps:	0

US-09-774-490-1 (1-2709) x ABG54764 (1-57)

QY 311 GAATGTGGAATCCACATGTGATCACTTCAATGGCTTGCCACAGCTCCAGTTAT 370
|||
Db 5 GluMetLeuGluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerIlyr 24

QY 371 CATACCTTCTTTGGATGAGGACGAGTACGCTGTATGTTGGCAAGATCACATA 430
|||
Db 25 HisThrPheLeuLeuAspGluGluArgSerArgLeuIlyValGlyAlaIlyAspHisIle 44

QY 431 TTTTCATTCGACCTGGTTAATATCAAGATTTTCAAAAG 469
|||
Db 45 PheSerPheAspLeuValIleIlyIlyAspPheGlnIly 57

RESULT 16

ABB39696
ID ABB39696 standard; Peptide; 57 AA.

XX AC ABB39696;

DT 04-FEB-2002 (first entry)

DE Peptide #7202 encoded by human foetal liver single exon probe.

KW Human; foetal liver; gene expression; single exon nucleic acid probe.

XX Homo sapiens.

XX WO200157277-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00669.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

DR WPI; 2001-483447/52.
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
PS Claim 27; SEQ ID NO 32331; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WFO at fwp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 57 AA;
XX
Alignment Scores:
Pred. No.: 1,23e-44 Length: 57
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0
DB: 22 Gaps: 0
XX
US-09-774-490-1 (1-2709) x ABB39696 (1-57)
XX
QY 311 GAATGTTGAATCCAAAGATGTGATCACTTTCAGTGGCCCAAGCTCCAGTTAT 370
DB 5 GllmetleuGluserAenValIlethrPheAenGlyLeuAlaAenSerSeryr 24
XX
QY 371 CATACCTTCCTTTTGATGAGAACGAGTAGCTGTATGTTGGCAAGATCACATA 430
DB 25 HlstrhPheLeuAenSpGluGluArgSerArgLeuYrValGlyAlaLysAspHisIle 44
XX
QY 431 TTTTCATTCGACCTGGTTAATATCAAGATTTTCAAAAG 469
DB 45 PheSerPheAspLeuValAenIleYrAspPheGlnLys 57
XX
RESULT 17
AAM60414
ID AAM60414 standard; Protein; 57 AA.
XX
AC AAM60414;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 32519.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 32519; 650pp + Sequence listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 57 AA;
XX
Alignment Scores:
Pred. No.: 1,23e-44 Length: 57
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0
DB: 22 Gaps: 0
XX
US-09-774-490-1 (1-2709) x AAM60414 (1-57)
XX
QY 311 GAATGTTGAATCCAAAGATGTGATCACTTTCAGTGGCCCAAGCTCCAGTTAT 370
DB 5 GllmetleuGluserAenValIlethrPheAenGlyLeuAlaAenSerSeryr 24
XX
QY 371 CATACCTTCCTTTTGATGAGAACGAGTAGCTGTATGTTGGCAAGATCACATA 430
DB 25 HlstrhPheLeuAenSpGluGluArgSerArgLeuYrValGlyAlaLysAspHisIle 44
XX
QY 431 TTTTCATTCGACCTGGTTAATATCAAGATTTTCAAAAG 469
DB 45 PheSerPheAspLeuValAenIleYrAspPheGlnLys 57
XX
RESULT 18
AAM73050
ID AAM73050 standard; Protein; 57 AA.
XX
AC AAM73050;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 33356.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; Leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX

XX WPI; 2001-488900/53.
DR Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
PT
XX
PS Example 4; SEQ ID NO: 33356; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 57 AA;
XX
Alignment Scores:
Pred. No.: 1,23e-44 Length: 57
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0
DB: 22 Gaps: 0
XX
US-09-774-490-1 (1-2709) x AAM73050 (1-57)
QY 311 GAAATGTTGGAATCCAAATGATGATCACTTTCAATGGCTTGCCCAAGCTCCAGTTAT 370
DB 5 GtmetleugluserashnshnvalletthrPheanglyleuAlasnserserTyr 24
QY 371 CATACCTTCCTTTGGATGAGGAACGAGTAGGCTGTATGTTGAGCAAGATCACATA 430
DB 25 HsttrPheleuleuAspGlucIuArgserArgleuTyrValGIyAlalysaphisile 44
QY 431 TTTTATTGACCTGCTGTATATATCAAGATTTTCAAAAG 469
DB 45 PheSerPheAspLeuValAsnIleLysAspPheGlnLys 57
XX
RESULT 19
AAM33272
ID AAM33272 standard; Protein; 57 AA.
XX
AC AAM33272;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #7309 encoded by probe for measuring placental gene expression.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-48897/53.
XX

XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
XX
PS Claim 27; SEQ ID No 33541; 654bp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP;
CC see A11315-A157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX
SQ Sequence 57 AA;
XX
Alignment Scores:
Pred. No.: 1,23e-44 Length: 57
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0
DB: 22 Gaps: 0
XX
US-09-774-490-1 (1-2709) x AAM33272 (1-57)
QY 311 GAAATGTTGGAATCCAAATGATGATCACTTTCAATGGCTTGCCCAAGCTCCAGTTAT 370
DB 5 GtmetleugluserashnshnvalletthrPheanglyleuAlasnserserTyr 24
QY 371 CATACCTTCCTTTGGATGAGGAACGAGTAGGCTGTATGTTGAGCAAGATCACATA 430
DB 25 HsttrPheleuleuAspGlucIuArgserArgleuTyrValGIyAlalysaphisile 44
QY 431 TTTTATTGACCTGCTGTATATATCAAGATTTTCAAAAG 469
DB 45 PheSerPheAspLeuValAsnIleLysAspPheGlnLys 57
XX
RESULT 20
ABG42894
ID ABG42894 standard; Peptide; 57 AA.
XX
AC ABG42894;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human peptide encoded by genome-derived single exon probe SEQ ID 32559.
XX
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Heremansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US00665.
XX
PR 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX

PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI, 2002-114183/15.
 XX
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 PS Claim 27, SEQ ID NO 32559; 634pp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemodiosclerosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein
 CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 57 AA;
 Alignment Scores:
 Pred. No.: 1,236-44 Length: 57
 Score: 53.00 Matches: 53
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 5.92% Indels: 0
 DB: 23 Gaps: 0
 US-09-774-490-1 (1-2709) x ABG42694 (1-57)
 QY 311 GAAATTTGGAATCCAAATGATGATCACTTTGATGGCTGGCCAAAGCTTCAAGTTAT 370
 DB 5 GUMERLEUGLUSERASNSNVALIETIRPHEASNGLYEULALANSERSERSERTYR 24
 QY 371 CATACCTCTCTTTGGATGAGAAAGAGAGAGCTGATGTTGATGAGCAAGATCAATCA 430
 DB 25 HLEHNPHELEUENUSPGUGLUGARGSERARGLERYVALJYALALYSASPHTSILE 44
 QY 431 TTTTCATTCGACCTGGTTAATATCAAGATTTTCAAG 469

DB 45 PHESERPHEASPEUVALSNILELYSASPENGLYS 57
 RESULT 21
 ID AAY21280
 AC AAY21280; Protein, 51 AA.
 XX
 XX 22-JUL-1999 (first entry)
 DT
 XX Human semaphorin III mutant protein fragment 13.
 DE
 XX Human, beta-amyloid precursor protein, beta-APP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGF-C; NSF-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 OS Synthetic.
 OS Homo sapiens.
 OS
 PN MO9845322-A2.
 PD 15-OCT-1998.
 XX
 PD 02-APR-1998; 98WO-1B00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 PR
 XX (UYOT-) RIKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACADEMY OF SCI.
 XX (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
 XX
 DR WPI, 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 PS Disclosure, Figure 16, 258pp; English.
 XX
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGF-C) and neuroendocrine specific protein A.
 CC
 SQ Sequence 51 AA;

PA (UYUT-) RIKSUNIV UTRECHT.
PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
PA (UYRO-) UNIV ROTTERDAM ERASMUS.
PI Burbach JPH, Grosveld FG, Van Leeuwen FW,
PI WPI; 1998-609901/51.
DR N-PSDB; AAX75767.
XX
XX
PT Diagnosing disease by detecting frameshift mutations in RNA or
PT corresponding protein mutations - used to diagnose cancer and
PT neurological diseases, particularly Alzheimer's disease, and also
PT for treatment and prevention with specific ribozymes or wild-type
PT RNA
PS Disclosure; Figure 16; 258bp; English.
XX
XX
CC This invention describes a novel method for the diagnosis of a disease
CC caused by, or associated with, an RNA molecule that has a frameshift
CC mutation. The method is used to diagnose age-related diseases, especially
CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
CC and many others listed) or susceptibility to these disorders. The method
CC allows a definitive diagnosis of Alzheimer's disease in living patients,
CC at an early stage. It is based on the observation that disease may be
CC caused by mutations in RNA rather than DNA. The invention describes the
CC use of neuronal system RNA molecules, specifically proteins including
CC beta-amyloid precursor protein (beta-APP), the microtubule associated
CC protein tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
CC protein-C (HMGP-C) and neuroendocrine specific protein A.
XX
XX
SQ Sequence 48 AA;
XX
XX
Alignment Scores:
Pred. No.: 1 62e-39 Length: 48
Score: 48.00 Matches: 48
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.36% Indels: 0
DB: 19 Gaps: 0
US-09-774-490-1 (1-2709) x AAY21324 (1-48)
QY 1393 TGTATTAACCTTGGCAAGAGTCCAGCCATGATCAATTCAGTGTTCATGAACA 1452
DB 1 CystylAsnleucylserylserSerHisValGlnserSerValSertYrGluGln 20
QY 1453 TGGCCCAATNGATCAAAAACGATGTAATTTCATTTACCAATTTCTGTACCG 1512
DB 21 SerProAsnSerAspGlnAsnGlyCysLysLeuSerIleTyrThaAsnCysArgArgPro 40
QY 1513 AGTGATGCGAAGATGACAGTA 1536
DB 41 SerGlyCysArgArgTrpThrVal 48
XX
XX
AC AAY21331;
XX
XX
DT 22-JUL-1999 (first entry)
DE Human semaphorin III mutant protein fragment 64.
XX
XX
XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
KW frameshift mutation; age-related disease; neurodegenerative disorder;
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;

KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KW neurofilament-F; presentin I; presentin II; cellular tumour antigen;
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
KW high mobility group protein-C; neuroendocrine specific protein A.
XX
XX
OS Synthetic.
OS Homo sapiens.
XX
XX
PN WO9845322-A2.
XX
XX
PD 15-OCT-1998.
XX
XX
PF 02-APR-1998; 98WO-1B00705.
XX
XX
PR 10-APR-1997; 97US-0043163.
XX
XX
PA (UYUT-) RIKSUNIV UTRECHT.
PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
PA (UYRO-) UNIV ROTTERDAM ERASMUS.
PI Burbach JPH, Grosveld FG, Van Leeuwen FW,
PI WPI; 1998-609901/51.
DR N-PSDB; AAX75767.
XX
XX
PT Diagnosing disease by detecting frameshift mutations in RNA or
PT corresponding protein mutations - used to diagnose cancer and
PT neurological diseases, particularly Alzheimer's disease, and also
PT for treatment and prevention with specific ribozymes or wild-type
PT RNA
XX
XX
PS Disclosure; Figure 16; 258bp; English.
XX
XX
XX
CC This invention describes a novel method for the diagnosis of a disease
CC caused by, or associated with, an RNA molecule that has a frameshift
CC mutation. The method is used to diagnose age-related diseases, especially
CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
CC and many others listed) or susceptibility to these disorders. The method
CC allows a definitive diagnosis of Alzheimer's disease in living patients,
CC at an early stage. It is based on the observation that disease may be
CC caused by mutations in RNA rather than DNA. The invention describes the
CC use of neuronal system RNA molecules, specifically proteins including
CC beta-amyloid precursor protein (beta-APP), the microtubule associated
CC protein tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
CC protein-C (HMGP-C) and neuroendocrine specific protein A.
XX
XX
SQ Sequence 46 AA;
XX
XX
Alignment Scores:
Pred. No.: 1 81e-37 Length: 46
Score: 46.00 Matches: 46
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.14% Indels: 0
DB: 19 Gaps: 0
US-09-774-490-1 (1-2709) x AAY21331 (1-46)
QY 1777 GTGTGGCCGCGCCGAGACCTTACTGCTGGAGTGGTTCGATGTCGCTATT 1836
DB 1 ValLeuProArgProArgProLeuLeuCysLeuGlyTrpPheCysMetPheSerLeuPhe 20
QY 1837 TCCCACTGCAAGAGACGCAAGACGCAAGATTAAGAAATGAGACCCACTGACTCA 1896

Db 21 SerHisCysLysGluThrHisIleThrArgTyrLysLysTyrPargProThrAspSer 40
QY 1897 CTGTCAGACTTACACCA 1914
|||||
Db 41 LeuPheArgLeuThrPro 46
RESULT 25
AAV21322
ID AAV21322 standard; Protein; 42 AA.
XX
XX AAV21322;
AC
XX 22-JUL-1999 (first entry)
DE Human semaphorin III mutant protein fragment 55.
KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
KW frameshift mutation; age-related disease; neurodegenerative disorder;
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
KW high mobility group protein-C; neuroendocrine specific protein A.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX WO9845322-A2.
XX
XX 15-OCT-1998.
XX
XX 02-APR-1998; 98WO-1B00705.
XX
XX 10-APR-1997; 97US-0043163.
XX
XX (UYUT-) RIJKSUNIV UTRECHT.
XX (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
XX (UYRO-) UNIV ROTTERDAM ERASMUS.
XX
XX Burbach JPH, Grosveld FG, Van Leeuwen FW;
PI
XX WPI; 1998-609901/51.
XX DR N-PSDB; AAX75767.
XX
XX Diagnosing disease by detecting frameshift mutations in RNA or
XX corresponding protein mutations - used to diagnose cancer and
XX neurological diseases, particularly Alzheimer's disease, and also
XX for treatment and prevention with specific ribozymes or wild-type
XX RNA
XX
XX Discloure; Figure 16; 258pp; English.
XX
XX This invention describes a novel method for the diagnosis of a disease
XX caused by, or associated with, an RNA molecule that has a frameshift
XX mutation. The method is used to diagnose age-related diseases, especially
XX cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
XX disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
XX multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
XX and many others listed) or susceptibility to these disorders. The method
XX allows a definitive diagnosis of Alzheimer's disease in living patients,
XX at an early stage. It is based on the observation that disease may be
XX caused by mutations in RNA rather than DNA. The invention describes the
XX use of neuronal system RNA molecules, specifically proteins including
XX beta-amyloid precursor protein (beta-APP), the microtubule associated
XX proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
XX associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
XX neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
XX protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
XX 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
XX protein-C (HMGP-C) and neuroendocrine specific protein A.

XX
XX SQ: Sequence 42 AA;
XX
XX Alignment Scores:
XX Score: 2,24e-33 Length: 42
XX Pred. No.: 42.00 Matches: 42
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 4.69% Indels: 0
XX DB: 19 Gaps: 0
XX
XX US-09-774-490-1 (1-2709) x AAV21322 (1-42)
QY 1240 TGTGAGAGGGTGTCTTGGTCATATGCCACAGAGGAGGCCAATCATATGGGT 1299
|||||
Db 1 CysGluYrGlyValProTyrPserIleCysProGlnGlyTyrThrGlnLeuSerMetCly 20
QY 1300 GCCTTATCAAGAGAGTCCCTATCCAGCGCCAGAACTGTCCACCAAACTTTGG 1359
|||||
Db 21 AlaLeuSerArgLysSerProLeuSerThrAlaArgAsnLeuSerGlnGlnIleTyr 40
QY 1360 TGGTTT 1365
|||||
Db 41 TrpPhe 42
XX
XX RESULT 26
XX ABG55921
XX ID ABG55921 standard; Peptide; 41 AA.
XX
XX ABG55921;
XX
XX 25-FEB-2003 (first entry)
XX
XX Human liver peptide, SEQ ID No 34569.
XX
XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
XX hypercholesterolaemia; coronary heart disease.
XX
XX Homo sapiens.
XX
XX WO200157273-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00664.
XX
XX 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488898/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analysing gene expression in human adult liver -
XX
XX Claim 27; SEQ ID No 34569; 658pp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (1) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 13109 defined nucleotide sequences given in the
XX specification (or complements/ fragments). The probe hybridises at high
XX stringency to a nucleic acid molecule expressed in the human adult
XX liver. (1) may be used for predicting, measuring and displaying gene
XX expression in samples derived from human adult liver. The genes
XX identified may be involved in genetic liver diseases such as cirrhosis,

QY 533 AAGAAATGCTTAATTCATCAAGTACTTAAGCATATATACAGACTGTTGACGCC 592
 DB 1 LysGluCysAlaAsnSphLeuValLeuValAlaTyrAsnGlnThrHisLeuTyrAla 20
 QY 593 TGTGGAACGGGGCTTTTCATCCAAATTTGCACCTTACATT 631
 DB 21 CysGlyThrGlyAlaPheHisProIleCysThrTyrIle 33
 RESULT 32
 ABB24410
 ID ABB24410 standard; Protein; 33 AA.
 AC ABB24410;
 XX
 XX
 DT 23-JAN-2002 (first entry)
 DE Protein #6409 encoded by probe for measuring heart cell gene expression.
 XX
 XX Human; gene expression; heart; microarray; vascular system;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 KM congenital heart disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200157274-A2.
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00666.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-488899/53.
 DR WPI; 2001-488899/53.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts -
 XX
 PS Claim 15; SEQ ID NO 26180; 530pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart (see
 CC ABA21533-ABA41305). The present sequence is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarrays.
 CC By measuring gene expression, the probes are useful for predicting, of the
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 33 AA;
 Alignment Scores:
 Pred. No.: 3.67e-24 Length: 33
 Score: 33.00 Matches: 33
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.69% Indels: 0
 DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x ABB24410 (1-33)
 QY 533 AAGAAATGCTTAATTCATCAAGTACTTAAGCATATATACAGACTGTTGACGCC 592
 DB 1 LysGluCysAlaAsnSphLeuValLeuValAlaTyrAsnGlnThrHisLeuTyrAla 20
 QY 593 TGTGGAACGGGGCTTTTCATCCAAATTTGCACCTTACATT 631
 DB 21 CysGlyThrGlyAlaPheHisProIleCysThrTyrIle 33
 RESULT 33
 AAM60513
 ID AAM60513 standard; Protein; 33 AA.
 AC AAM60513;
 XX
 XX
 DT 05-NOV-2001 (first entry)
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 32618.
 XX
 XX Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KM epilepsy; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200157275-A2.
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00667.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-483446/52.
 DR WPI; 2001-483446/52.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT brains -
 XX
 PS Example 4; SEQ ID NO: 32618; 650pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancer. The present sequence is a protein encoded by one of
 CC the probes of the invention.
 XX
 SQ Sequence 33 AA;
 Alignment Scores:
 Pred. No.: 3.67e-24 Length: 33
 Score: 33.00 Matches: 33
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.69% Indels: 0
 DB: 22 Gaps: 0
 US-09-774-490-1 (1-2709) x AAM60513 (1-33)
 QY 533 AAGAAATGCTTAATTCATCAAGTACTTAAGCATATATACAGACTGTTGACGCC 592
 DB 1 LysGluCysAlaAsnSphLeuValLeuValAlaTyrAsnGlnThrHisLeuTyrAla 20
 QY 593 TGTGGAACGGGGCTTTTCATCCAAATTTGCACCTTACATT 631
 DB 21 CysGlyThrGlyAlaPheHisProIleCysThrTyrIle 33

Db 1 LysGluCysAlaAsnPhelIeLysValIleuLysAlaTyrAsnGlnThrHisLeuTyrAla 20
 QY 593 TGTGAAACGGGGGCTTTTCATTCAGAGTAAAGGATATATATGAGCTGCTGTAAGCC 592
 DB 21 CysGlyThrGlyAlaPheHisProIleCysThrTyrIle 33

RESULT 34

AAM73166
 ID AAM73166 standard; Protein; 33 AA.

AC AAM73166;
 DT 06-NOV-2001 (first entry)

XX Human bone marrow expressed probe encoded protein SEQ ID NO: 33472.

DE Human bone marrow expressed exon; gene expression analysis; probe;
 XX microarray; cancer; leukemia; lymphoma; myeloma.

KW Homo sapiens.
 OS Homo sapiens.

XX WO200157276-A2.
 PN 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00668.
 PF 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.
 DR 2001-488900/53.

XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -

XX Example 4; SEQ ID NO: 33472; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention.

XX SQ Sequence 33 AA;
 SQ

Alignment Scores:
 Pred. No.: 3, 67e-24 Length: 33
 Score: 33.00 Matches: 33
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.69% Indels: 0
 DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x AAM73166 (1-33)

QY 533 AAGAAATGTGCTAATTTCATTCAGAGTAAAGGATATATATGAGCTGCTGTAAGCC 592
 DB 1 LysGluCysAlaAsnPhelIeLysValIleuLysAlaTyrAsnGlnThrHisLeuTyrAla 20

QY 593 TGTGAAACGGGGGCTTTTCATTCAGAGTAAAGGATATATATGAGCTGCTGTAAGCC 631
 DB 21 CysGlyThrGlyAlaPheHisProIleCysThrTyrIle 33

RESULT 35
 AAM33379
 ID AAM33379 standard; Protein; 33 AA.
 XX AAM33379;
 AC AAM33379;
 DT 17-OCT-2001 (first entry)

XX Peptide #7416 encoded by probe for measuring placental gene expression.
 DE Probe; microarray; human; placenta; antenatal diagnosis;
 XX genetic disorder.

XX Homo sapiens.
 OS Homo sapiens.

XX WO200157272-A2.
 PN 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00663.
 PF 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488897/53.
 DR 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -

XX Claim 27; SEQ ID NO 33648; 654bp; English.

XX The present invention relates to single exon nucleic acid probes (SENPs;
 CC see A1131315-A1151546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders.

XX SQ Sequence 33 AA;
 SQ

Alignment Scores:
 Pred. No.: 3, 67e-24 Length: 33
 Score: 33.00 Matches: 33
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.69% Indels: 0
 DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x AAM33379 (1-33)

QY 533 AAGAAATGTGCTAATTTCATTCAGAGTAAAGGATATATATGAGCTGCTGTAAGCC 592
 DB 1 LysGluCysAlaAsnPhelIeLysValIleuLysAlaTyrAsnGlnThrHisLeuTyrAla 20

QY 593 TGTGAAACGGGGGCTTTTCATTCAGAGTAAAGGATATATATGAGCTGCTGTAAGCC 631
 DB 21 CysGlyThrGlyAlaPheHisProIleCysThrTyrIle 33

RESULT 36
 ABG43013
 ID ABG43013 standard; Peptide; 33 AA.
 AC ABG43013;
 XX

19-AUG-2002 (first entry)
Human peptide encoded by genome-derived single exon probe SEQ ID 32678.
Human; single exon probe; asthma; lung cancer; COPD; ILD;
chronic obstructive pulmonary disease; interstitial lung disease;
familial idiopathic pulmonary fibrosis; neurofibromatosis;
tuberosclerosis; Gaucher's disease; Niemann-Pick disease;
Hernandez-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
pulmonary histiocytosis; lymphangioleiomyomatosis; Kargener syndrome;
pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
primary ciliary dyskinesia; pulmonary hypertension;
hyaline membrane disease.
Homo sapiens.
WO200186003-A2.
15-NOV-2001.
30-JAN-2001; 2001WO-US00665.
04-FEB-2000; 2000US-180312P.
26-MAY-2000; 2000US-207456P.
30-JUN-2000; 2000US-0608408.
03-AUG-2000; 2000US-0632366.
21-SEP-2000; 2000US-234687P.
27-SEP-2000; 2000US-236359P.
04-OCT-2000; 2000GB-0024263.
(MOLE-) MOLECULAR DYNAMICS INC.
Penn SG, Hanzel DK, Chen W, Rank DR;
WPI; 2002-114183/15.
Spatially-addressable set of single exon nucleic acid probes, used to
measure gene expression in human lung samples -
Claim 27; SEQ ID NO 32678; 634bp; English.
The invention relates to a spatially-addressable set of single exon
nucleic acid probes for measuring gene expression in a sample derived
from human lung comprising single exon nucleic acid probes having one of
12614 nucleic acid sequences mentioned in the specification, or their
complements or the 12387 open reading frames derived from the 12614
probes. Also included are a microarray comprising the novel set of
probes; the novel set of probes which hybridise at high stringency to a
nucleic acid expressed in the human lung; measuring gene expression in a
sample derived from human lung, comprising (a) contacting the array with
a collection of detectably labeled nucleic acids derived from human lung
mRNA, and (b) measuring the label detectably bound to each probe of
the array; identifying exons in a eukaryotic genome, comprising
(a) algorithmically predicting at least one exon from genomic sequences
of the eukaryote; and (b) detecting specific hybridisation of detectably
labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
having a fragment identical to the predicted exon, the probe is included
in the above mentioned microarray; assigning exons to a single gene,
comprising (a) identifying exons from genomic sequence by the method
above and (b) measuring the expression of each of the exons in several
tissues and/or cell types using hybridisation to a single exon
microarrays having a probe with the exon, where a common pattern of
expression of the exons in the tissues and/or cell types indicates that
the exons should be assigned to a single gene; a peptide comprising one
of 12011 sequences, mentioned in the specification, or encoded by the
probes/open reading frames (ORF). The probes are used for gene
expression analysis, and for identifying exons in a gene, particularly
using human lung derived mRNA and for the study of lung diseases
such as asthma, lung cancer, chronic obstructive pulmonary disease
(COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
Niemann-Pick disease, Hernandez-Pudlak syndrome, sarcoidosis, pulmonary
haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,

pulmonary alveolar proteinosis, Kargener syndrome, fibrocystic
pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
and hyaline membrane disease. The present sequence is a peptide/protein
encoded by a single exon probe of the invention.
CC Note: The sequence data for this patent did not form part
of the printed specification, but was obtained in electronic
format directly from WIPO at
ftp://wipo.int/pub/published_pct_sequences.
SQ Sequence 33 AA;
Alignment Scores:
Pred. No.: 3,67e-24 Length: 33
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3,69# Indels: 0
DB: Gaps: 0
US-09-774-490-1 (1-2709) x ABG43013 (1-33)
QY 533 AAAGATGCTGCTAATTTTCATCAGGTACTTAAGCATATATCAGACTCTTGACGCC 592
DB 1 LysglLysAlaAsnPhelIleValLeuYsaIaTyraSngGlnThrHisLeuTyAla 20
QY 593 TGTGGACGGGGGCTTTTCATCCATTCGACCTACACT 631
DB 21 CysglYThrGlyAlaPheHisProIleCysThrTyIle 33
RESULT 37
AAV21292
ID AAV21292 standard; Protein; 31 AA.
XX
XX AAV21292;
AC
XX 22-JUN-1999 (first entry)
DT
XX
XX Human semaphorin III mutant protein fragment 25.
DE
XX
XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
KX frameshift mutation; age-related disease; neurodegenerative disorder;
KX Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
KX Huntington's disease; multiple sclerosis; alcoholic liver disease;
KX diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KX ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KX neurofilament-F; presenilin 1; presenilin II; cellular tumour antigen;
KX glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
KX bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
KX high mobility group protein-C; neuroendocrine specific protein A.
OS
OS Synthetic.
OS Homo sapiens.
XX
XX MO9845322-A2.
XX
XX 15-OCT-1998.
PD
XX 02-APR-1998; 98WO-1B00705.
PF
XX 10-APR-1997; 97US-0043163.
PR
XX (UVT-) RIJKSUNIV UTRECHT.
PA (ROYA-) ROYAL NETHERLANDS ACADEMY OF SCI.
PA (UTRO-) UNIV ROTTERDAM ERASMUS.
XX
XX Burbach JPH, Grosveld FG, Van Leeuwen FW;
PI WPI; 1998-609901/51.
DR N-PSDB; AAX75767.
XX
XX Diagnosing disease by detecting frameshift mutations in RNA or
PT corresponding protein mutations - used to diagnose cancer and
PT neurological diseases, particularly Alzheimer's disease, and also

PT for treatment and prevention with specific ribozymes or wild-type
 RNA

PS Disclosure, Figure 16; 258bp; English.

CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease, II
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-ApP), the microtubule associated
 CC proteins tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGp-C) and neuroendocrine specific protein A.

CC Sequence 31 AA;

Alignment Scores:

Pred. No.:	4,09e-22	Length:	31
Score:	31.00	Matches:	31
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.46%	Indels:	0
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) x AAY21292 (1-31)

QY 2190 CCTGTGAGTCAATTCACAGACATTTGAGAGACTTCTTCAATAAGATGATGAG 2249

DB 1 ProliplySerleuthrInserleleTplyAaenPheleleylethetnetwetu 20

QY 2250 ATGGCTCTAAGACCAAGAAATGTCATTCGCA 2282

DB 21 MetAlaLeuArgProlylsyCysProIleAla 31

RESULT 38

AAY21279 standard; Protein, 30 AA.

ID AAY21279

AC AAY21279;

XX 22-JUL-1999 (first entry)

DE Human semaphorin III mutant protein fragment 12.

XX Human; beta-amyloid precursor protein; beta-ApP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presentin I; presentin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGp-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.

OS Synthetic.

OS Homo.sapiens.

XX MO9845322-A2.

XX 15-OCT-1998.

PD 02-APR-1998; 98WO-IB00705.

XX 10-APR-1997; 97US-0043163.

XX (UYUT-) RIKSUNIV UTRECHT.

XX (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.

XX (UYRO-) UNIV ROTTERDAM ERASMUS.

XX Butbach JPH, Grosveld FG, Van Leeuwen FW;

XX WPI, 1998-609901/51.

XX N-PSDB; AAX75767.

XX Diagnosing disease by detecting frameshift mutations in RNA or

XX corresponding protein mutations - used to diagnose cancer and

XX neurological diseases, particularly Alzheimer's disease, and also

XX for treatment and prevention with specific ribozymes or wild-type

XX RNA

XX Disclosure, Figure 16; 258bp; English.

CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-ApP), the microtubule associated
 CC proteins tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGp-C) and neuroendocrine specific protein A.

CC Sequence 30 AA;

Pred. No.:	4.33e-21	Length:	30
Score:	30.00	Matches:	30
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.35%	Indels:	0
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) x AAY21279 (1-30)

QY 1143 ACTTTAAGATCCCTAAATCCAGTGTATGAGTGTATACGATCCAGTACATTT 1202

DB 1 ThrleuUySileleUySileGlnleuUyIwctGluCysleuArgleuProvalThrIhe 20

QY 1203 TCAAGGATCAGCCGCTGTATGATACGA 1232

DB 21 SerArgapsInProCysValCysIleAla 30

RESULT 39

AAY21269 standard; Protein, 27 AA.

ID AAY21269

AC AAY21269;

XX 22-JUL-1999 (first entry)

DE Human semaphorin III mutant protein fragment 2.

XX Human; beta-amyloid precursor protein; beta-ApP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;

KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN MO9845322-A2.
 XX
 PD 15-OCT-1998.
 XX
 PF 02-APR-1998; 98WO-IB00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIJKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
 DR WPI, 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 PS Disclosure; Figure 16; 258pp; English.
 XX
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-ApP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.
 XX
 SO Sequence 27 AA;
 Alignment Scores:
 Pred. No.: 5,11e-18 Length: 27
 Score: 27.00 Matches: 27
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 3.02% Indels: 0
 DB: 19 Gaps: 0
 US-09-774-490-1 (1-2709) x AAY21269 (1-27)
 QY 213 CTAGAGTTGTCTGTCTTTCTGGAGATTAATTACAGCAAGCAAACTATCAGAATG 272
 Db 1 LeuGlyLeuSerValPheSerGlyGlyIuTyrrTyrluGlnGlnGlnGlnThrIleArgMet 20
 QY 273 GGAAGAACAAATGTGCCAAGGC 293
 Db 21 GlyArgThrMetCysGlnGly 27

RESULT 40
 AAY21296
 ID AAY21296 standard; Protein; 26 AA.
 XX
 AC AAY21296;
 XX
 DT 22-JUL-1999 (first entry)
 XX
 DE Human semaphorin III mutant protein fragment 29.
 XX
 KW Human, beta-amyloid precursor protein; beta-ApP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN MO9845322-A2.
 XX
 PD 15-OCT-1998.
 XX
 PF 02-APR-1998; 98WO-IB00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIJKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Burbach JPH, Grosveld FG, Van Leeuwen FW;
 DR WPI, 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 PS Disclosure; Figure 16; 258pp; English.
 XX
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-ApP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.
 XX
 SO Sequence 26 AA;
 Alignment Scores:
 Pred. No.: 5.41e-17 Length: 26

Score: 26.00 Matches: 26
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.91% Indels: 0
DB: 19 Gaps: 0

US-09-774-490-1 (1-2709) x AAY21296 (1-26)

QY 2595 ACCTTTTTCATGCGCATTTATGATGATGCGGAATTCAGCTGAGTTCACC 2654
|||||
1 ThrPhePheMetAlaLeuGlyCysLeuGlnTrpIleGlnLeuSerSerThr 20
DB 2655 AATTATTAATTAATCA 2672
|||||
21 AsnTyrLysLeuAsnPro 26

RESULT 41
AAY21328
ID AAY21328 standard; Protein; 25 AA.

XX AAY21328;
AC
XX
XX 22-JUL-1999 (first entry)
DT
XX
XX Human semaphorin III mutant protein fragment 61.
DE
XX
XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
KW frameshift mutation; age-related disease; neurodegenerative disorder;
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KW neurofilament-E; presenilin I; presenilin II; cellular tumour antigen;
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
KW high mobility group protein-C; neuroendocrine specific protein A.
XX
XX
OS Synthetic.
OS Homo sapiens.
XX
XX W09845322-A2.
XX
XX 15-OCT-1998.
XX
XX PD 02-APR-1998; 98MO-IB00705.
XX
XX PF 10-APR-1997; 97US-0043163.
XX
XX PR
XX (UYUT-) RIKSUNIV UTRECHT.
XX (ROYA-) ROYAL NETHERLANDS ACADEMIES OF SCI.
XX (UYRO-) UNIV ROTTERDAM ERASMUS.
XX
XX PA
XX Burbach JPH, Grosveld FG, Van Leeuwen FW;
XX
XX PI
XX WPI; 1998-609901/51.
XX
XX DR N-PSDB; AAX75767.
XX
XX PT Diagnosing disease by detecting frameshift mutations in RNA or
XX corresponding protein mutations - used to diagnose cancer and
XX neurological diseases, particularly Alzheimer's disease, and also
XX for treatment and prevention with specific ribozymes or wild-type
XX RNA
XX
XX PS Disclosure; Figure 16; 258bp; English.
XX
XX This invention describes a novel method for the diagnosis of a disease
XX caused by, or associated with, an RNA molecule that has a frameshift
XX mutation. The method is used to diagnose age-related diseases, especially
XX cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
XX disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
XX multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
XX and many others listed) or susceptibility to these disorders. The method
XX allows a definitive diagnosis of Alzheimer's disease in living patients,

CC at an early stage. It is based on the observation that disease may be
CC caused by mutations in RNA rather than DNA. The invention describes the
CC use of neuronal system RNA molecules, specifically proteins including
CC beta-amyloid precursor protein (beta-APP), the microtubule associated
CC protein tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
CC protein-C (HMGP-C) and neuroendocrine specific protein A.
XX
XX SQ Sequence 25 AA;
XX

Alignment Scores:
Pred. No.: 5 726-16 Length: 25
Score: 25.00 Matches: 25
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.79% Indels: 0
DB: 19 Gaps: 0

US-09-774-490-1 (1-2709) x AAY21328 (1-25)

QY 1612 TTGAGAGGCTTCTGCTGGAGAAATGACACTTTTCGGAGCCGACTGCTATTTCAGC 1671
|||||
1 PheArgArgGlySerAlaGlyArgAsnAspSerPheSerGlyThrAspCysTyrPheSer 20
DB 1672 AATGAGCTTTCCAC 1686
|||||
21 AsnGlyAlaPheHis 25

RESULT 42
AAY21285
ID AAY21285 standard; Protein; 24 AA.

XX AAY21285;
AC
XX
XX 22-JUL-1999 (first entry)
DT
XX
XX Human semaphorin III mutant protein fragment 18.
DE
XX
XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
KW frameshift mutation; age-related disease; neurodegenerative disorder;
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
KW high mobility group protein-C; neuroendocrine specific protein A.
XX
XX
OS Synthetic.
OS Homo sapiens.
XX
XX W09845322-A2.
XX
XX 15-OCT-1998.
XX
XX PD 02-APR-1998; 98MO-IB00705.
XX
XX PF 10-APR-1997; 97US-0043163.
XX
XX PR
XX (UYUT-) RIKSUNIV UTRECHT.
XX (ROYA-) ROYAL NETHERLANDS ACADEMIES OF SCI.
XX (UYRO-) UNIV ROTTERDAM ERASMUS.
XX
XX PA
XX Burbach JPH, Grosveld FG, Van Leeuwen FW;
XX
XX PI
XX WPI; 1998-609901/51.
XX
XX DR N-PSDB; AAX75767.
XX
XX PT Diagnosing disease by detecting frameshift mutations in RNA or

PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA

PS Disclosure; Figure 16; 258pp; English.

CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.

CC Sequence 24 AA;

Alignment Scores:

Pred. No.:	6,05e-15	Length:	24
Score:	24.00	Matches:	24
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.68%	Indels:	0
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) x AAY21285 (1-24)

QY 1509 ACCGATGAGTGCAGATGAGTATGATGTTATCGGAACAGATGTTGGA 1568

DB 1 ThrgutrtPmGlnLysMecAspserMecLysCysLeuSerGlnGlnMetLeuGly 20

QY 1569 CCGTCTTAAAG 1580

DB 21 ProPheLeuLys 24

RESULT 43

ID AAY21337 standard; Protein; 23 AA.

AC AAY21337;

DT 22-JUL-1999 (first entry)

DE Human semaphorin III mutant protein fragment 70.

KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presentin I; presentin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.

OS Synthetic.

OS Homo sapiens.

XX WO9845322-A2.

XX 15-OCT-1998.

XX 02-APR-1998; 98WO-1B00705.
 XX 10-APR-1997; 97US-0043163.

PA (UYUT-) RIJKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACADEMIES & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.

PI Burbach JPH, Grosveld FG, Van Leeuwen FW;

DR WPI; 1998-609901/51.
 DR N-PSDB; AAX75767.

PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA

PS Disclosure; Figure 16; 258pp; English.

CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.

CC Sequence 23 AA;

Alignment Scores:

Pred. No.:	6.4e-14	Length:	23
Score:	23.00	Matches:	23
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.57%	Indels:	0
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) x AAY21337 (1-23)

QY 2113 TCTACACAGAGATTCAGCAATTCCTGCGATGCGTGGAGATGGGTTGATCA 2172

DB 1 SerThrThrgLgUgLyPhaArgGlnLeuProLeuProCysGlyGlyThrTrpValHisThr 20

QY 2173 AACTTTTCT 2181

DB 21 AsnSerSer 23

RESULT 44

ID AAY21310 standard; Protein; 23 AA.

AC AAY21310;

DT 22-JUL-1999 (first entry)

DE Human semaphorin III mutant protein fragment 43.

KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;

KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9845322-A2.
 XX
 PD 15-OCT-1998.
 XX
 XX 02-APR-1998; 98WO-IB00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Burdach JPH, Grosveld FG, Van Leeuwen FW;
 XX
 DR WPI: 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 PS Disclosure; Figure 16; 258pp; English.
 XX
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-ApP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.
 CC
 XX
 SQ Sequence 23 AA;
 Alignment Scores:
 Pred. No.: 6, 4e-14 Length: 23
 Score: 23.00 Matches: 23
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2, 57% Indels: 0
 DB: 19 Gaps: 0
 US-09-774-490-1 (1-2709) x AAY21310 (1-23)
 OY 778 TTTTGGGGGAGAGCTTGTATCTTCGAACTCTTGGGACCAACCAATCAGAGAC 837
 DB 1 PheTYRGlyAlaArgLeucyStryLeuProAsnSerTrpAlaProProAsnGlnAsp 20
 OY 838 AGAGCAGCA 846

DB |||||
 21 ArgAlaAla 23
 RESULT 45
 AAY21290.
 ID AAY21290 standard; Protein; 23 AA.
 AC AAY21290;
 XX
 DT 22-JUL-1999 (first entry)
 XX
 DE Human semaphorin III mutant protein fragment 23.
 XX
 KW Human; beta-amyloid precursor protein; beta-ApP; diagnosis; cancer;
 KW frameshift mutation; age-related diseases; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9845322-A2.
 XX
 PD 15-OCT-1998.
 XX
 XX 02-APR-1998; 98WO-IB00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Burdach JPH, Grosveld FG, Van Leeuwen FW;
 XX
 DR WPI: 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 PS Disclosure; Figure 16; 258pp; English.
 XX
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-ApP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.
 CC
 XX
 SQ Sequence 23 AA;

CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.

XX Sequence 67 AA;

XX Alignment Scores:

Score: 5.86e-13 Length: 67
 Percent Similarity: 22.00 Matches: 35
 Best Local Similarity: 97.22% Conservative: 0
 Query Match: 2.46% Mismatches: 0
 Indels: 1
 Gaps: 0

US-09-774-490-1 (1-2709) x AAG02551 (1-67)

QY 176 AAGAGGACCTACGCGCTTCGACGATGGCTGTTACTAGATTTGCTCTTCTCGG 235

Db 33 TTTGGTThyThyTserValCysSerMetGlyTyrPLeuThrArgIleValCysLeuPheTyr 52

QY 236 GGAGTATTACTTACAGCAAGCAACTATCAGATGGAGACA 281

Db 53 -GluTyrTyrLeuGlnGlnGlnGlnThrIleArgMetGlyArgThr 67

RESULT 48

AAV21342 AAV21342 standard; Protein; 21 AA.

XX AAV21342;

XX 22-JUL-1999 (first entry)

XX Human semaphorin III mutant protein fragment 75.

XX Human, beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presentin I; presentin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.

XX Synthetic.

XX Homo sapiens.

XX W09845322-A2.

XX 15-OCT-1998.

XX 02-APR-1998; 98WO-1B00705.

XX 10-APR-1997; 97US-0043163.

XX (UYUT-) RIJKSUNIV UTRECHT.

XX (ROYA-) ROYAL NETHERLANDS ACADEMY OF SCI.

XX (UTRO-) UNIV ROTTERDAM ERASMUS.

XX Burbach JPH, Grosveld FG, Van Leeuwen FW;

XX MPI; 1998-609901/51.

XX N-PSDB; AAX75767.

XX Diagnosing disease by detecting frameshift mutations in RNA or

PT corresponding protein mutations - used to diagnose cancer and

PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA

XX Disclosure; Figure 16; 258pp; English.

XX This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.

XX Sequence 21 AA;

XX Alignment Scores:

Score: 7.17e-12 Length: 21
 Percent Similarity: 21.00 Matches: 21
 Best Local Similarity: 100.00% Conservative: 0
 Query Match: 2.35% Mismatches: 0
 Indels: 0
 Gaps: 0

US-09-774-490-1 (1-2709) x AAV21342 (1-21)

QY 2293 CCAGAGGCTGTTGTTACAGAGCACTTCATGACCTGATCAACCCCAATCTCAACAGCAT 2352

Db 1 ProGluGlyLeuValGlnArgLeuHisAlaHisGlnProGlnSerGlnHisasp 20

QY 2353 GGA 2355

Db 21 Gly 21

RESULT 49

AAV21343 AAV21343 standard; Protein; 21 AA.

XX AAV21343;

XX 22-JUL-1999 (first entry)

XX Human semaphorin III mutant protein fragment 76.

XX Human, beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presentin I; presentin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.

XX Synthetic.

XX Homo sapiens.

XX W09845322-A2.

XX 15-OCT-1998.

XX


```

PF 02-APR-1998; 98WO-1B00705.
XX
XX 10-APR-1997; 97US-0043163.
XX
XX (UYUT-) RIJKSUNIV UTRECHT.
PA (ROYA-) ROYAL NETHERLANDS ACADEMY OF SCI.
PA (UYRO-) UNIV ROTTERDAM ERASMUS.
XX
XX Bubach JPH, Grosveld FG, Van Leeuwen FW;
XX WPI; 1998-609901/51.
XX N-PSDB; AAX75767.
XX
XX Diagnosing disease by detecting frameshift mutations in RNA or
XX corresponding protein mutations - used to diagnose cancer and
XX neurological diseases, particularly Alzheimer's disease, and also
XX for treatment and prevention with specific ribozymes or wild-type
XX RNA
XX
XX Disclosure; Figure 16; 258pp; English.
XX
XX This invention describes a novel method for the diagnosis of a disease
XX caused by, or associated with, an RNA molecule that has a frameshift
XX mutation. The method is used to diagnose age-related diseases, especially
XX cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
XX disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
XX multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
XX and many others listed) or susceptibility to these disorders. The method
XX allows a definitive diagnosis of Alzheimer's disease in living patients,
XX at an early stage. It is based on the observation that disease may be
XX caused by mutations in RNA rather than DNA. The invention describes the
XX use of neuronal system RNA molecules, specifically proteins including
XX beta-amyloid precursor protein (beta-Ap), the microtubule associated
XX proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
XX associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
XX neurofilament-F, presentin I, presentin II, glial fibrillary acidic
XX protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
XX 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
XX protein-C (HMGP-C) and neuroendocrine specific protein A.
XX
XX SQ Sequence 21 AA;
XX
XX Alignment Scores:
XX Pred. No.: 7,17e-12 Length: 21
XX Score: 21.00 Matches: 21
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 2,35# Indels: 0
XX DB: 19 Gaps: 0
XX
XX US-09-774-490-1 (1-2709) x AAY21343 (1-21)
XX
XX QY 2368 ACAAGTTTGAAAGAGGACCGAAGCAAGTCGCGAAGGCGAGACATACCCGAGGAA 2427
XX |||||
XX 1 ThrsrLrLgUlyGlyProlyThrThrserAlaLysArgThrTrpArgGlu 20
XX
XX QY 2428 CAG 2430
XX |||||
XX 21 Gln 21
XX
XX Db
XX
XX RESULT 50
XX AAY21278
XX ID AAY21278 standard; Protein; 21 AA.
XX
XX AC AAY21278;
XX
XX 22-JUL-1999 (first entry)
XX
XX DE Human semaphorin III mutant protein fragment 11.
XX
XX Human, beta-amyloid precursor protein; beta-Ap; diagnosis; cancer;
XX frameshift mutation; age-related disease; neurodegenerative disorder;
XX Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;

```

```

XX Huntington's disease; multiple sclerosis; alcoholic liver disease;
XX diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
XX ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
XX neurofilament-F; presentin I; presentin II; cellular tumour antigen;
XX glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
XX bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
XX high mobility group protein-C; neuroendocrine specific protein A.
XX
XX OS Synthetic.
XX
XX OS Homo sapiens.
XX
XX EN WO9845322-A2.
XX
XX 15-OCT-1998.
XX
XX 02-APR-1998; 98WO-1B00705.
XX
XX 10-APR-1997; 97US-0043163.
XX
XX (UYUT-) RIJKSUNIV UTRECHT.
XX (ROYA-) ROYAL NETHERLANDS ACADEMY OF SCI.
XX (UYRO-) UNIV ROTTERDAM ERASMUS.
XX
XX Bubach JPH, Grosveld FG, Van Leeuwen FW;
XX WPI; 1998-609901/51.
XX N-PSDB; AAX75767.
XX
XX Diagnosing disease by detecting frameshift mutations in RNA or
XX corresponding protein mutations - used to diagnose cancer and
XX neurological diseases, particularly Alzheimer's disease, and also
XX for treatment and prevention with specific ribozymes or wild-type
XX RNA
XX
XX Disclosure; Figure 16; 258pp; English.
XX
XX This invention describes a novel method for the diagnosis of a disease
XX caused by, or associated with, an RNA molecule that has a frameshift
XX mutation. The method is used to diagnose age-related diseases, especially
XX cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
XX disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
XX multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
XX and many others listed) or susceptibility to these disorders. The method
XX allows a definitive diagnosis of Alzheimer's disease in living patients,
XX at an early stage. It is based on the observation that disease may be
XX caused by mutations in RNA rather than DNA. The invention describes the
XX use of neuronal system RNA molecules, specifically proteins including
XX beta-amyloid precursor protein (beta-Ap), the microtubule associated
XX proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
XX associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
XX neurofilament-F, presentin I, presentin II, glial fibrillary acidic
XX protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
XX 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
XX protein-C (HMGP-C) and neuroendocrine specific protein A.
XX
XX SQ Sequence 21 AA;
XX
XX Alignment Scores:
XX Pred. No.: 7,17e-12 Length: 21
XX Score: 21.00 Matches: 21
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 2,35# Indels: 0
XX DB: 19 Gaps: 0
XX
XX US-09-774-490-1 (1-2709) x AAY21278 (1-21)
XX
XX QY 1074 TTTGCTCAGTCCAGGTCGCAATGCGACATTTTGATGACGAGATGAT 1133
XX |||||
XX 1 PheAlaGlnCysGlnValGlnMetAlaLeuThrLeuLeuMetAsnGlyArgMetCys 20
XX
XX QY 1134 TCC 1136
XX |||||

```

Db	21	Ser	21
RESULT	51		
AA21329			
ID	AA21329	standard; Protein; 20 AA.	
XX			
XX			
AC	AA21329;		
XX			
XX			
DT	22-UTL-1999	(first entry)	
XX			
DE	Human semaphorin III mutant protein fragment 62.		
KW	Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;		
KW	frameshift mutation; age-related disease; neurodegenerative disorder;		
KW	Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;		
KW	Huntington's disease; multiple sclerosis; alcoholic liver disease;		
KW	diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;		
KW	ubiquitin B, apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;		
KW	neurofilament-F; prenenilin I; prenenilin II; cellular tumour antigen;		
KW	glial fibrillary acidic protein; GRAP; p53; semaphorin III; HUPF-1;		
KW	dcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;		
KW	high mobility group protein-C; neuroendocrine specific protein A.		
XX			
XX			
OS	Synthetic.		
OS	Homo sapiens.		
XX			
PN	WO945322-A2.		
XX			
PD	15-OCT-1998.		
XX			
PF	02-APR-1998; 98WO-IB00705.		
XX			
PR	10-APR-1997; 97US-0043163.		
XX			
XX	(UYUT-) RIJKSUNIV UTRECHT.		
PA	(ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.		
PA	(UYRO-) UNIV ROTTERDAM ERASMUS.		
XX			
PL	Burbach JPH, Grosveld FG, Van Leeuwen FW;		
XX			
XX	WPI, 1998-609901/51.		
DR	N-PSDB; AAX75767.		
XX			
PT	Diagnosing disease by detecting frameshift mutations in RNA or		
PT	corresponding protein mutations - used to diagnose cancer and		
PT	neurological diseases, particularly Alzheimer's disease, and also		
PT	for treatment and prevention with specific ribozymes or wild-type		
PT	RNA		
XX			
XX			
XX	Disclosure; Figure 16; 258pp; English.		
XX			
XX	This invention describes a novel method for the diagnosis of a disease		
CC	caused by, or associated with, an RNA molecule that has a frameshift		
CC	mutation. The method is used to diagnose age-related diseases, especially		
CC	cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's		
CC	disease, Down's syndrome, myotonic dystrophy, Huntington's disease,		
CC	multiple sclerosis, alcoholic liver disease, diabetes mellitus type II		
CC	and many others listed) or susceptibility to these disorders. The method		
CC	allows a definitive diagnosis of Alzheimer's disease in living patients,		
CC	at an early stage. It is based on the observation that disease may be		
CC	caused by mutations in RNA rather than DNA. The invention describes the		
CC	use of neuronal system RNA molecules, specifically proteins including the		
CC	beta-amyloid precursor protein (beta-APP), the microtubule associated		
CC	proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule		
CC	associated protein 2 (MAP2), neurofilament-L, neurofilament-M,		
CC	neurofilament-F, prenenilin I, prenenilin II, glial fibrillary acidic		
CC	protein (GRAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma		
CC	2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group		
CC	protein-C (HMGP-C) and neuroendocrine specific protein A.		
XX			
XX			
Sequence	20 AA;		
Alignment Scores:			

Pred. No.:	7.6e-11	Length:	20
Score:	20.00	Matches:	20
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.23%	Indels:	0
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) x AAY21329 (1-20)

OY 1690 GCAGCACACTATATATTGGTTCAACGCGCTGGGCTTCCAGCTCCTTTACACCGGTG 1749
Db 1 AlAlatHrThrlleTyTYTrPheAsnGLyTyrGLyCysProAlaProPheThrProVal 20
RESULT 52
AAY21318
ID AAY21318 standard; Protein, 19 AA.
XX
AC AAY21318;
XX
DT 22-JUL-1999 (first entry)
XX
DE Human semaphorin III mutant protein fragment 51.
XX
KW Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
KW frameshift mutation; age-related disease; neurodegenerative disorder;
KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
KW neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;
KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HMPF-1;
KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; NMSP-C; NSP-A;
KW high mobility group protein-C; neuroendocrine specific protein A.
XX
OS Synthetic.
XX
OS Homo sapiens.

WO9845322-A2.
XX
PN 15-OCT-1998.
XX
PD 02-APR-1998; 98MO-IB00705.
XX
PF 10-APR-1997; 97US-0043163.
XX
PR (OYUT-) RIJXSUNIV UTRECHT.
PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
PA (UYRO-) UNIV ROTTERDAM ERASMUS.
FI Burbach JPH, Grosveld FG, Van Leeuwen FW;
PI WPI; 1998-609901/51.
DX N-PSDB; AAX75767.
XX
DR Diagnosing disease by detecting frameshift mutations in RNA or
PT corresponding protein mutations - used to diagnose cancer and
PT neurological diseases, particularly Alzheimer's disease, and also
PT for treatment and prevention with specific ribozymes or wild-type
PT RNA
XX
PS Disclosure; Figure 16; 258bp; English.

This invention describes a novel method for the diagnosis of a disease caused by, or associated with, an RNA molecule that has a frameshift mutation. The method is used to diagnose age-related diseases, especially cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's disease, Down's syndrome, myotonic dystrophy, Huntington's disease, multiple sclerosis, alcoholic liver disease, diabetes mellitus type II and many others listed) or susceptibility to these disorders. The method allows a definitive diagnosis of Alzheimer's disease in living patients, at an early stage. It is based on the observation that disease may be caused by mutations in RNA rather than DNA. The invention describes the used of neuronal system RNA molecules, specifically proteins including

DR WPI; 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and also
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 PS Disclosure; Figure 16; 258pp; English.
 XX
 XX This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease, CC
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-ApP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGp-C) and neuroendocrine specific protein A.
 XX
 XX Sequence 18 AA;
 SQ
 Alignment Scores:
 Pred. No.: 8.53e-09 Length: 18
 Score: 18.00 Matches: 18
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 2.01% Indels: 0
 DB: 19 Gaps: 0
 US-09-774-490-1 (1-2709) x AAY21309 (1-18)
 QY 721 GCTGCTGACGACATCCCTTTTATAGATGAGAAATTACTGTGAACTGCAGC 774
 Db 1 AAlaAlaSpSerIleProPheAnaArgTlPargIleIleLeuTrpAnCySer 18
 RESULT 55
 AAY21301
 ID AAY21301 standard; Protein; 17 AA.
 XX
 AC AAY21301;
 XX
 DT 22-JUL-1999 (first entry)
 XX
 DE Human semaphorin III mutant protein fragment 34.
 XX
 KW Human; beta-amyloid precursor protein; beta-ApP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presentin I; presentin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGp-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9845322-A2.
 XX
 PD 15-OCT-1998.
 XX

XX
 PF 02-APR-1998; 98WO-1B00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIJKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UYRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Butbach JPH, Grosveld FG, Van Leeuwen FW;
 XX
 DR WPI; 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 PS Disclosure; Figure 16; 258pp; English.
 XX
 XX This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease, CC
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-ApP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGp-C) and neuroendocrine specific protein A.
 XX
 XX Sequence 17 AA;
 SQ
 Alignment Scores:
 Pred. No.: 9.04e-08 Length: 17
 Score: 17.00 Matches: 17
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.90% Indels: 0
 DB: 19 Gaps: 0
 US-09-774-490-1 (1-2709) x AAY21301 (1-17)
 QY 451 TATCAAGATTTTCAAAAGATTGTGCGCAGATCTTACACCGAGAGAGA 501
 Db 1 TyrGInGlyPheSerIyASpCyValAlaSerIleLeuHISGInLyArg 17
 RESULT 56
 AAG62702
 ID AAG62702 standard; peptide; 17 AA.
 XX
 AC AAG62702;
 XX
 DT 17-SEP-2001 (first entry)
 XX
 DE Hanatoxin-like sequence from type 3 semaphorin Sema3A.
 XX
 KW Hanatoxin; tarentula; toxin; voltage-gated potassium channel;
 KW voltage-gated calcium channel; hanatoxin-like sequence; HTS;
 KW semaphorin; dorsal root ganglion repulsion; growth cone collapse.
 XX
 OS Mammalia.
 XX

PN WO200138491-A2.
 XX 31-MAY-2001.
 XX 07-NOV-2000; 2000WO-US41943.
 XX 08-NOV-1999; 99US-0164056.
 XX (GENO) GEN HOSPITAL CORP.
 XX Behar O, Woolf CJ;
 XX WPI; 2001-451494/48.
 DR
 XX
 XX Polypeptide sequences that encompass the hanatoxin-like sequences of
 PT semaphorin, useful as a drugs to treat any condition or disease that
 PT is characterized by abnormal calcium channel function -
 XX
 XX Claim 3; Page 10; 29pp; English.
 XX
 XX AAG62702-22 represent hanatoxin-like sequences (HTLS) found in
 CC the semaphorin domain of mammalian secreted semaphorins. Hanatoxin
 CC is a tarantula toxin that selectively blocks some voltage-gated
 CC potassium and calcium channels. The HTLS is responsible for the dorsal
 CC root ganglion repulsion and growth cone collapse activities associated
 CC with semaphorins. Polypeptides containing HTLS can be used to modulate
 CC the activity of calcium channels. The peptides can also be used as an
 CC antigen to generate antibodies that can then be used to modulate the
 CC activity of calcium channels by inactivating naturally occurring channel
 CC ligands. The peptides or antibodies can be used as drugs to treat any
 CC condition or disease that is characterized by abnormal calcium channel
 CC function.
 CC
 XX
 XX Sequence 17 AA:
 SQ
 Alignment Scores:
 Pred. No.: 9,04e-08 Length: 17
 Score: 17.00 Matches: 17
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.90% Indels: 0
 DB: Gaps: 0
 US-09-774-490-1 (1-2709) x AAG62702 (1-17)
 QY 1763 AAGGTCGTGAGTGTGCTGCGCCGAGACCTTACTGCTTGAGAT 1813
 Db 1 AAGGTCGTGAGTGTGCTGCGCCGAGACCTTACTGCTTGAGAT 17
 1 AAGGTCGTGAGTGTGCTGCGCCGAGACCTTACTGCTTGAGAT 17
 RESULT 57
 AAB84219
 ID AAB84219 standard; Protein; 779 AA.
 XX
 XX AAB84219;
 AC
 XX
 XX 06-AUG-2001 (first entry)
 DT
 XX
 XX Amino acid sequence of a semaphorin polypeptide designated ZSMF-16.
 DE
 XX
 XX Human; semaphorin; ZSMF-16; neurite growth; neurite outgrowth;
 KW T lymphocyte suppressor; cancer; stroke; brain damage; paralysis;
 KW spinal injury; neurodegenerative disease; amyotrophic lateral sclerosis;
 KW Alzheimer's disease; Huntington's disease; Parkinson's disease;
 KW peripheral neuropathy; demyelinating disease; multiple sclerosis;
 KW immunosuppression; autoimmune disease; insulin dependent diabetes;
 KW rheumatoid arthritis.
 KW
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Peptide 1..22
 FT Protein 23..779
 FT note= "signal peptide"

PT Modified-site /note= "mature protein"
 FT 62
 FT /note= "N-glycosylation site"
 FT 76..500
 FT Domain
 FT /note= "semaphorin"
 FT Modified-site 124
 FT /note= "N-glycosylation site"
 FT 593..654
 FT Domain
 FT /note= "Ig-like domain"
 FT Modified-site 594
 FT /note= "N-glycosylation site"
 FT
 PN WO200140278-A2.
 XX
 XX 07-JUN-2001.
 PD
 XX
 XX 06-DEC-2000; 2000WO-US31116.
 PF
 XX
 XX 06-DEC-1999; 99US-0455560.
 PR
 XX
 XX (ZYMO) ZYMOGENETICS INC.
 PA
 XX
 XX Holloway JL, Foley KP;
 PI
 XX
 XX WPI; 2001-374784/39.
 DR
 XX
 XX N-PSDB; AAF90250.
 DR
 XX
 XX Novel human semaphorin polypeptide, ZSMF-16, useful for treating
 PT peripheral neuropathies Alzheimer's and Huntington's disease and
 PT polynucleotide encoding ZSMF-16 useful for detecting genetic
 PT abnormality and cancer -
 XX
 XX
 XX Claim 13; Page 121-123; 124pp; English.
 PS
 XX
 XX The present sequence represents a semaphorin polypeptide, designated
 CC ZSMF-16. ZSMF-16 is a neurite growth and development modulator. It also
 CC enhances spinal cord and sensory neurite outgrowth and patterning, and
 CC is involved in the activation and regulation of T lymphocytes suppressor.
 CC ZSMF-16 is useful for detecting a genetic abnormality or cancer. ZSMF-16
 CC polynucleotide probes can be used to detect 3p21 loss, trisomy,
 CC duplication or translocation associated with mammary tumor tissue,
 CC breast tumour, liver, small intestine, bone cancers, etc.. ZSMF-16 can
 CC be used to modulate neurite growth and development and demarcate nervous
 CC system structures. ZSMF-16 are also useful for regenerating and directing
 CC neurite outgrowths following strokes, brain damage caused by head
 CC injuries, paralysis caused by spinal injuries, and for treating
 CC neurodegenerative diseases such as amyotrophic lateral sclerosis,
 CC Alzheimer's disease, Huntington's disease, Parkinson's disease and
 CC peripheral neuropathies, or demyelinating diseases e.g., multiple
 CC sclerosis. ZSMF-16 also acts as a mediator of immunosuppression,
 CC and thus useful for diagnosing and treating autoimmune diseases such as
 CC insulin dependent diabetes, rheumatoid arthritis, and multiple
 CC sclerosis. It can also be used as an antiinflammatory for inhibition
 CC of antigen in humoral and cellular immunity and for immunosuppression
 CC in graft and organ transplants.
 CC
 XX
 XX Sequence 779 AA:
 SQ
 Alignment Scores:
 Pred. No.: 5.51e-08 Length: 779
 Score: 17.00 Matches: 17
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.90% Indels: 0
 DB: Gaps: 0
 US-09-774-490-1 (1-2709) x AAB84219 (1-779)
 QY 1766 GCGTGTGCTGAGTGTGCTGCGCCGAGACCTTACTGCTTGAGAT 1816
 Db 525 AAGGTCGTGAGTGTGCTGCGCCGAGACCTTACTGCTTGAGAT 541
 RESULT 58

AA078481	AA078481 standard; Protein; 779 AA.
ID	AA078481
AC	AA078481;
XX	
XX	12-APR-2002 (first entry)
DT	
XX	Human ZSMF-16.
DE	
XX	Semaphorin; cancer; immune response; ZSMF-16; human;
KW	neuroprotective; nootropic; multiple sclerosis;
KW	Alzheimer's disease; Parkinson's disease; stroke.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Peptide
FT	1..22
FT	/label= signal_peptide
FT	23..779
FT	/label= mature_protein
FT	23..75
FT	/note= "N-terminal region"
FT	62
FT	/note= "N- glycosylated"
FT	76..500
FT	/label= semaphorin_domain
FT	124
FT	/note= "N- glycosylated"
FT	501..592
FT	/label= Middle_domain
FT	593..654
FT	/label= Ig-like_domain
FT	594
FT	/note= "N- glycosylated"
FT	655..738
FT	/label= C-terminal_domain
FT	744..751
FT	/note= "ATP/GTP binding site motif"
XX	
XX	US2001049432-A1.
PN	
XX	
PD	06-DEC-2001.
XX	
XX	06-DEC-2000; 2000US-0731179.
PF	
XX	
PR	06-DEC-1999; 99US-169238P.
XX	
PA	(HOLL/) HOLLOWAY J L.
XX	(FOLEY/) FOLEY K P.
XX	
PI	Holloway JL, Foley KP;
XX	
DR	WPI: 2002-121401/16.
DR	N-PSDB; ABA97362.
XX	
PT	Polynucleotides encoding a semaphorin polypeptide expressed in breast
PT	cancer and neuronal tissue, useful for treating neurological damage and
PT	disease, and to detect cancer -
PS	Claim 1; Page 33-35; 59pp; English.
XX	
CC	This invention relates to an isolated polynucleotide encoding a
CC	semaphorin polypeptide expressed in breast cancer and neuronal tissue.
CC	The polypeptide is neuroprotective and nootropic in its action.
CC	The semaphorin polypeptide can be used to modulate neurite growth
CC	and development, for example to treat peripheral neuropathies, stroke,
CC	brain damage, from brain and spine injuries, and neurodegenerative
CC	diseases such as multiple sclerosis, Alzheimer's and Parkinson's
CC	disease. It can also be used to modulate cellular activation,
CC	targeting, adhesion, proliferation and differentiation and
CC	immunological responses. Molecules of the invention can also be used to
CC	treat cancer. This sequence represents the ZSMF-16 which is a member of
CC	the semaphorin family.

XX	Sequence	779 AA;
SQ	Alignment Scores:	
Pred. NO.:	5.51e-08	Length: 779
Score:	17.00	Matches: 17
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	1.90%	Indels: 0
Gaps:	23	Gaps: 0
US-09-774-490-1 (1-2709) x AAG78481 (1-779)		
OY	1766 GCCTGTCGTGAGTGGTTGCCTCGCCGAGACCCCTTACTGTGCTTGGAATGCT	1816
Db	525 AlAcyslaadlucscyspseudalaaagaprotlyrcysalatrpaspgly	541
RESULT 59		
AAB23609		
ID	AAB23609 standard; Protein; 782 AA.	
XX		
AC	AAB23609;	
XX		
DT	12-JAN-2001 (first entry)	
XX		
DE	Human secreted protein SEQ ID NO: 18.	
XX		
KW	Human; secreted protein; cytokine; cell proliferation; nutritional supplement; immune modulation; autoimmune disorder; haematopoiesis regulation; tissue growth; haemostasis; inflammation.	
OS	Homo sapiens.	
XX		
FH	Key Location/Qualifiers	
FT	Peptide 10..22	
FT	/label= signal_peptide	
FT	Protein 23..782	
FT	/label= mature_protein	
PN	WC0200049134-A1.	
PD	24-AUG-2000.	
PF	18-FEB-2000; 2000MO-US04340.	
PR	19-FEB-1999; 99US-0120680.	
PR	23-APR-1999; 99US-0298733.	
PR	17-AUG-1999; 99US-0149639.	
PR	23-SEP-1999; 99US-0155686.	
PR	01-OCT-1999; 99US-0157247.	
PR	29-NOV-1999; 99US-0167822.	
PR	29-NOV-1999; 99US-0167823.	
PR	15-FEB-2000; 2000US-0298733.	
PA	(ALPH-) ALPHAGEN INC.	
PI	Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;	
DR	WPI; 2000-549267/50.	
XX	N-PDBE; AAA93109.	
PT	New secreted proteins and polynucleotides encoding them, which are derived from Homo sapiens, useful for therapy, diagnosis, and research, as well as nutritional sources or supplements -	
PS	Claim 27, Page 250-253; 309pp; English.	
XX	The present sequence is the sequence of a human secreted protein. Its cDNA was isolated from an adult brain cDNA library. The proteins and coding sequences of the invention can be used in the isolation of similar genes and proteins, in the elucidation of their function in vivo, and to treat a number of conditions. It is possible that they may have uses as nutritional supplements, as cytokine or cell proliferation	

CC factors, in immune modulation, where they may be used to treat immune and
CC autoimmune diseases, as haematopoiesis regulators (treating myeloid or
CC lymphoid cell deficiencies), in the promotion of tissue growth, they may
CC have chemokine or chemotactic activity, haemostatic or thrombolytic
CC activity, or anti-inflammatory activity.

XX SQ Sequence 782 AA;

Alignment Scores:

Pred. No.:	5.5e-08	Length:	782
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.90%	Indels:	0
DB:	21	Gaps:	0

US-09-774-490-1 (1-2709) x AAB23609 (1-782)

OY 1766 GCGTGTGCTGAGTGTGCTGCGCCGAGACCTTACTGTGCTGGAGTGT 1816

DB 528 AAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaIleTrrpAspGly 544

RESULT 60

ID AAG65620 standard; Protein: 782 AA.

XX AAG65620;

XX 07-JAN-2002 (first entry)

XX Novel human protein (NHP) sequence.

XX NHP; novel human protein; secreted protein; semaphorin; oxytocin;

XX neurohypophyseal; neotropic; gene therapy; drug screening.

XX Homo sapiens.

XX WO200170806-A2.

XX 27-SEP-2001.

XX 20-MAR-2001; 2001WO-US088834.

XX 20-MAR-2000; 2000US-190638P.

XX 22-MAR-2000; 2000US-191188P.

XX 31-MAR-2000; 2000US-193639P.

XX (LEXI-) LEXICON GENETICS INC.

XX Walke DW, Wilganowski NL, Turner CA, Hilbun E, Wang X, Donoho G;

XX Scoville J;

XX WPI, 2001-611483/70.

XX N-PSDB; AA47791.

XX New polynucleotides encoding human proteins that share structural

XX similarity with semaphorin proteins, protein hormones of

XX neurohypophyseal family for drug screening, diagnosis and therapy of

XX biological disorders

XX Claim 4; Page 38-40; 43pp; English.

XX The invention relates to novel human secreted proteins (NHP) that share
XX structural similarity with semaphorin proteins, protein/peptide hormones
XX of the neurohypophyseal family and oxytocin (neurohypophysis I precursor)
XX family. The NHP nucleotide sequences are useful in drug screening
XX techniques for treating symptomatic or phenotypic manifestations of
XX perturbing the normal function of NHP in the body. Nucleotide constructs
XX encoding NHP products are useful in gene therapy for modulating NHP
XX expression. The constructs can be used to genetically engineer host cells
XX to express NHP products in vivo, these genetically engineered cells
XX function as bioreactors in the body delivering a continuous supply of a
XX NHP, a NHP peptide, or a NHP fusion protein to the body. The sequences

CC also find use in molecular mutagenesis/evolution of proteins that are
CC partially encoded by the NHP sequences. The encoded NHP polypeptides are
CC useful for generating antibodies, as reagents in diagnostic assays, for
CC identifying other cellular gene products related to NHP and as reagents
CC in assays for screening for compounds that are useful in the treatment of
CC mental, biological or medical disorders and diseases. The present
CC sequence represents the amino acid sequence of a NHP.

XX SQ Sequence 782 AA;

Alignment Scores:

Pred. No.:	5.5e-08	Length:	782
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.90%	Indels:	0
DB:	22	Gaps:	0

US-09-774-490-1 (1-2709) x AAG65620 (1-782)

OY 1766 GCGTGTGCTGAGTGTGCTGCGCCGAGACCTTACTGTGCTGGAGTGT 1816

DB 528 AAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaIleTrrpAspGly 544

RESULT 61

ID AAB23636 standard; Protein: 785 AA.

XX AAB23636;

XX 12-JAN-2001 (first entry)

XX Human secreted protein SEQ ID NO: 92.

XX Human secreted protein; cytokine; cell proliferation;

XX nutritional supplement; immune modulation; autoimmune disorder;

XX haematopoiesis regulation; tissue growth; haemostasis; inflammation.

XX Homo sapiens.

XX WO200049134-A1.

XX 24-AUG-2000.

XX 18-FEB-2000; 2000WO-US04340.

XX 19-FEB-1999; 99US-0120680.

XX 23-APR-1999; 99US-0298733.

XX 17-AUG-1999; 99US-0149639.

XX 23-SEP-1999; 99US-0155686.

XX 01-OCT-1999; 99US-0157247.

XX 29-NOV-1999; 99US-0167822.

XX 29-NOV-1999; 99US-0167823.

XX 15-FEB-2000; 2000US-0298733.

XX (ALPH-) ALPHAGENE INC.

XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;

XX WPI, 2000-549267/50.

XX New secreted proteins and polynucleotides encoding them, which are
XX derived from Homo sapiens, useful for therapy, diagnosis, and research,
XX as well as nutritional sources or supplements -
XX Disclosure; Page 298-300; 309pp; English.
XX The present invention is concerned with a number of secreted proteins
XX and their coding sequences isolated from various human cDNA libraries.
XX The proteins and coding sequences can be used in the isolation of
XX similar genes and proteins, in the elucidation of their function in vivo,
XX and to treat a number of conditions. It is possible that they may have
XX uses as nutritional supplements, as cytokine or cell proliferation

CC factors, in immune modulation, where they may be used to treat immune and
CC autoimmune diseases, as haematopoiesis regulators (treating myeloid or
CC lymphoid cell deficiencies), in the promotion of tissue growth, they may
CC have chemokine or chemotactic activity, haemostatic or thrombolytic
CC activity, or anti-inflammatory activity. No information about sequences
CC AAB2632-B2645 is given in the specification.

XX Sequence 785 AA;

Alignment Scores:

Pred. No.:	5.5e-08	Length:	785
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.90%	Indels:	0
DB:	21	Gaps:	0

US-09-774-490-1 (1-2709) x AAB2636 (1-785)

QY 1766 GCGTGTGCTGAGTGTTCCTGCGCCGAGACCTTACTGTGCTGGATGCT 1816

Db 531 AAlaCySAAlaGlucyScySAleuAlaArgAspProTyrCySAIaTrpAspGly 547

RESULT 62

AAG65619

ID AAG65619 standard; Protein; 875 AA.

AC AAG65619;

DT 07-JAN-2002 (first entry)

DE Novel human protein (NHP) sequence.

KW NHP; novel human protein; secreted protein; semaphorin; oxytocin;

KW neurohypophyseal; neurotropic; gene therapy; drug screening.

OS Homo sapiens.

PN WO200170806-A2.

PD 27-SEP-2001.

PF 20-MAR-2001; 2001WO-US08834.

PR 20-MAR-2000; 2000US-190638P.

PR 22-MAR-2000; 2000US-191188P.

PR 31-MAR-2000; 2000US-193639P.

PA (LEXI-) LEXICON GENETICS INC.

PI Walke DW, Wilganowski NL, Turner CA, Hilbun E, Wang X, Donoho G,

PI Scoville J;

DR WPI; 2001-611483/70.

DR N-PSDB; AA447790.

PT New polynucleotides encoding human proteins that share structural

PT similarity with semaphorin proteins, protein hormones of

PT neurohypophyseal family for drug screening, diagnosis and therapy of

PT biological disorders -

PS Claim 2; Page 35-37; 43pp; English.

XX The invention relates to novel human secreted proteins (NHP) that share
XX structural similarity with semaphorin proteins, protein/peptide hormones
XX of the neurohypophyseal family and oxytocin (neurohypophysin 1 precursor)
XX family. The NHP nucleotide sequences are useful in drug screening
XX techniques for treating symptomatic or phenotypic manifestations of
XX perturbing the normal function of NHP in the body. Nucleotide constructs
XX encoding NHP products are useful in gene therapy for modulating NHP
XX expression. The constructs can be used to genetically engineer host cells
XX to express NHP products in vivo, these genetically engineered cells
XX function as bioreactors in the body delivering a continuous supply of a

CC NHP, a NHP peptide, or a NHP fusion protein to the body. The sequences
CC also find use in molecular mutagenesis/evolution of proteins that are
CC partially encoded by the NHP sequences. The encoded NHP polypeptides are
CC useful for generating antibodies, as reagents in diagnostic assays, for
CC identifying other cellular gene products related to NHP and as reagents
CC in assays for screening for compounds that are useful in the treatment of
CC mental, biological or medical disorders and diseases. The present
CC sequence represents the amino acid sequence of a NHP.

XX Sequence 875 AA;

Alignment Scores:

Pred. No.:	5.42e-08	Length:	875
Score:	17.00	Matches:	17
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.90%	Indels:	0
DB:	22	Gaps:	0

US-09-774-490-1 (1-2709) x AAG65619 (1-875)

QY 1766 GCGTGTGCTGAGTGTTCCTGCGCCGAGACCTTACTGTGCTGGATGCT 1816

Db 621 AAlaCySAAlaGlucyScySAleuAlaArgAspProTyrCySAIaTrpAspGly 637

RESULT 63

AAG62704

ID AAG62704 standard; peptide; 17 AA.

AC AAG62704;

DT 17-SEP-2001 (first entry)

DE Hanatoxin-like sequence from type 3 semaphorin msemab3B.

KW Hanatoxin; tarantula; toxin; voltage-gated potassium channel;

KW Voltage-gated calcium channel; hanatoxin-like sequence; HTLS;

KW semaphorin; dorsal root ganglion repulsion; growth cone collapse.

OS Mus sp.

PN WO200138491-A2.

PD 31-MAY-2001.

PF 07-NOV-2000; 2000WO-US41943.

PF 08-NOV-1999; 99US-0164056.

PA (GEHO) GEN HOSPITAL CORP.

PI Behar O, Woolf CJ;

PI WPI; 2001-451494/48.

PT Polypeptide sequences that encompass the hanatoxin-like sequences of

PT semaphorins, useful as a drug to treat any condition or disease that

PT is characterized by abnormal calcium channel function -

PS Claim 3; Page 10; 29pp; English.

XX AAG62702-22 represent hanatoxin-like sequences (HTLS) found in
XX the semaphorin domain of mammalian secreted semaphorins. Hanatoxin
XX is a tarantula toxin that selectively blocks some voltage-gated
XX potassium and calcium channels. The HTLS is responsible for the dorsal
XX root ganglion repulsion and growth cone collapse activities associated
XX with semaphorins. Polypeptides containing HTLS can be used to modulate
XX the activity of calcium channels. The peptides can also be used as an
XX antigen to generate antibodies that can then be used to modulate the
XX activity of calcium channels by inactivating naturally occurring channel
XX ligands. The peptides or antibodies can be used as drugs to treat any
XX condition or disease that is characterized by abnormal calcium channel
XX function.

XX SQ Sequence 17 AA;
 Alignment Scores:
 Pred. No.: 9.52e-07 Length: 17
 Score: 16.00 Matches: 16
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.79% Indels: 0
 DB: 22 Gaps: 0
 US-09-774-490-1 (1-2709) x AAG62704 (1-17)
 OY 1766 GCGTGTGAGAGTGTGCTCGCCGAGACCTTACTGCTGGAT 1813
 DB 2 AAlcYsAlaGlUCyCysLeuAlaArgAProTyrCysAlaTrpAsp 17
 RESULT 64
 AAY21300
 ID AAY21300 standard; Protein; 15 AA.
 AC AAY21300;
 XX 22-JUL-1999 (first entry)
 DT Human semaphorin III mutant protein fragment 33.
 DE Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 XX frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presentin I; presentin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9845322-A2.
 XX
 PD 15-OCT-1998.
 XX
 PF 02-APR-1998; 98WO-IB00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIJKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UTRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Butbach JPH, Grosveld FG, Van Leeuwen FW;
 PI WPI; 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 PT RNA
 XX
 PS Disclosure; Figure 16; 258bp; English.
 XX
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method

CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.
 XX
 SQ Sequence 15 AA;
 Alignment Scores:
 Pred. No.: 1.02e-05 Length: 15
 Score: 15.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.68% Indels: 0
 DB: 19 Gaps: 0
 US-09-774-490-1 (1-2709) x AAY21300 (1-15)
 OY 403 GCTGTATGTGAGCAAGATCACATTTTCATTCGACCTGGT 447
 DB 1 AAlaValCysTrpSerIleGlySerHisIleIlehelAeArgProGly 15
 RESULT 65
 AAY21281
 ID AAY21281 standard; Protein; 15 AA.
 AC AAY21281;
 XX 22-JUL-1999 (first entry)
 DT Human semaphorin III mutant protein fragment 14.
 DE Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 XX frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presentin I; presentin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9845322-A2.
 XX
 PD 15-OCT-1998.
 XX
 PF 02-APR-1998; 98WO-IB00705.
 XX
 PR 10-APR-1997; 97US-0043163.
 XX
 PA (UYUT-) RIJKSUNIV UTRECHT.
 PA (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.
 PA (UTRO-) UNIV ROTTERDAM ERASMUS.
 XX
 PI Butbach JPH, Grosveld FG, Van Leeuwen FW;
 PI WPI; 1998-609901/51.
 DR N-PSDB; AAX75767.
 XX
 PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type

PT RNA
 XX
 PS Disclosure; Figure 16; 258pp; English.
 CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC protein tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presenilin I, presenilin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HMPF-1, high mobility group
 CC protein-C (HMGPC) and neuroendocrine specific protein A.
 XX
 SQ Sequence 15 AA;

Alignment Scores:
 Pred. No.: 1,02e-05 Length: 15
 Score: 15.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.68% Indels: 0
 DB: 19 Gaps: 0

US-09-774-490-1 (1-2709) x AAY21281 (1-15)

QY 1401 CCTTGGCAAGAGTCATCCAGCCATGATCCAGCTGTTCTTA 1445

DB 1 ProLeuGInGluValIleGlnProCysThrIleGlnCysPheLeu 15

RESULT 66

AA62727 standard; peptide; 749 AA.

AC AA62727;

DT 17-SEP-2001 (first entry)

DE Amino acid sequence of human semaphorin Sema3B.

KW Hanatoxin; tarentula; toxin; voltage-gated potassium channel;

KW voltage-gated calcium channel; hanatoxin-like sequence; HTLS;

KW semaphorin; dorsal root ganglion repulsion; growth cone collapse.

OS Homo sapiens.

PN WO200138491-A2.

PD 31-MAY-2001.

PF 07-NOV-2000; 2000WO-US41943.

PR 08-NOV-1999; 99US-0164056.

PA (GEHO) GEN HOSPITAL CORP.

PI Behar O, Woolf CJ;

DR WPI; 2001-451494/48.

PT Polypeptide sequences that encompass the hanatoxin-like sequences of

PT semaphorin, useful as a drugs to treat any condition or disease that

PT is characterized by abnormal calcium channel function -

PS Claim 6; Page 12; 29pp; English.

XX
 CC The present sequence represents a semaphorin. The specification
 CC describes hanatoxin-like sequences (HTLS) found in the semaphorin
 CC domain of mammalian secreted semaphorins. Hanatoxin is a tarentula
 CC toxin that selectively blocks some voltage-gated potassium and calcium
 CC channels. The HTLS is responsible for the dorsal root ganglion repulsion
 CC and growth cone collapse activities associated with semaphorins.
 CC Polypeptides containing HTLS can be used to modulate the activity of
 CC calcium channels. The peptides can also be used as an antigen to
 CC generate antibodies that can then be used to modulate the activity of
 CC calcium channels by inactivating naturally occurring channel ligands.
 CC The peptides or antibodies can be used as drugs to treat any condition
 CC or disease that is characterized by abnormal calcium channel function.

XX
 SQ Sequence 749 AA;

Alignment Scores:
 Pred. No.: 6.13e-06 Length: 749
 Score: 15.00 Matches: 15
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.68% Indels: 0
 DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x AA62727 (1-749)

QY 1028 AGAAGTCGCGGATTAATGACACATTCCTCAAGCTGCTTG 1072

DB 277 ArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeu 291

RESULT 67

AA30617 standard; Protein; 751 AA.

AC AA30617;

DT 31-MAR-1999 (first entry)

DE Human semaphorin E protein from clone BR5334.

KW Human; semaphorin E; clone BR5334; nutritional; immune stimulating;

KW vaccine; haematopoiesis regulating activity; tissue growth;

KW haemostatic; thrombolytic; anti-inflammatory; tumour inhibition;

OS Homo sapiens.

PN WO9853065-A1.

PD 26-NOV-1998.

PF 19-MAY-1998; 98WO-US10188.

PR 18-MAY-1998; 98US-0080695.

PR 19-MAY-1997; 97US-0858834.

PA (GENETICS INST INC.

PI Agostino MJ, Fechtel K, Howes SH, Jacobs K, Lavalie ER;

PI McCoy JM, Werberg D, Racie LA, Spaulding V, Treacy M;

DR WPI; 1999-059742/05.

DR N-PSDB; AAX03792.

PT New polypeptides encoding secreted human proteins - derived from a

PT human foetal kidney cDNA library

PS Claim 7; Page 44-47; 58pp; English.

CC The present sequence is human semaphorin E from cDNA clone BR5334.

CC Human semaphorin E polynucleotide sequences and protein sequences from

CC the present invention, are predicted to have biological activities which

CC would make them suitable for treating, preventing or ameliorating
CC medical conditions in humans and animals, although no supporting data is
CC given. Suggested activities include nutritional activity, immune
CC stimulating (e.g. as vaccines) or suppressing activity, haematopoiesis
CC regulating activity, tissue growth activity, activin/inhibin activity,
CC chemotactic/chemokinetic activity, haemostatic and thrombolytic
CC activity, receptor/ligand activity, anti-inflammatory activity,
CC cadherin/tumour invasion suppressor activity, and tumour inhibition
CC activity. The polynucleotide sequences are also stated to be useful for
CC gene therapy. A host cell transfected with the polynucleotide sequence
CC for recombinant production of the clone BR533_4 related protein.

SQ Sequence 751 AA;

Alignment Scores:

Pred. No.:	6,13e-06	Length:	751
Score:	15.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.68%	Indels:	0
DB:	20	Gaps:	0

US-09-774-490-1 (1-2709) x AAM30617 (1-751)

OY 1028 AGAAGCTGTGTAATTAATGACAACTTCCTCAAGCTGCTG 1072

DB 274 ArgSerLeuValAsnLysTrpThrRhrPheLeuLysAlaArgLeu 288

RESULT 68

ID AAB28379 standard; Protein; 751 AA.

XX AAB28379;

XX 19-FEB-2001 (first entry)

XX Clone BR533_4.

XX Human; clone BR533_4; secreted protein; immune deficiency;
XX microbial infection; autoimmune disorder; allergy; asthma; inflammation;
XX myeloid deficiency; lymphoid cell deficiency; anaemia; burn;
XX wound healing; ulcer; periodontal disease; nervous system disease;
XX neuropathy; lung fibrosis; liver fibrosis; cancer.

XX Homo sapiens.

XX WO200063692-A1.

XX 26-OCT-2000.

XX 14-APR-2000; 2000MO-US10048.

XX 15-APR-1999; 99US-0292550.

XX (GENY) GENETICS INST INC.

XX Jacobs K, McCoy JM, Lavallie ER, Collins-Racie LA, Evans C;

XX Metberg D, Treacy W, Spaulding V,

XX WPI; 2000-679620/66.

XX N-PSDB; AAC68800.

XX New monoclonal antibodies, useful for treating cancer and
XX immunodetection of secreted proteins which are in turn useful for
XX treating neurological, inflammatory, immune diseases and microbial
XX infections -

XX Claim 8; Pages 67-69; 75pp; English.

XX The present sequence is human clone BR533_4. The coding sequence for the
XX present sequence was isolated from a human foetal kidney cDNA library.
XX The BR533_4 protein is a secreted protein and can be used to treat a

CC number of conditions including various immune deficiencies and disorders,
CC microbial infections, autoimmune disorders, allergic reactions such as
CC asthma, respiratory problems, inflammation, myeloid or lymphoid cell
CC deficiencies, anaemias, burns, wound healing, ulcers, periodontal
CC disease, central and peripheral nervous system diseases and neuropathies,
CC lung or liver fibrosis and cancer.

SQ Sequence 751 AA;

Alignment Scores:

Pred. No.:	6,13e-06	Length:	751
Score:	15.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.68%	Indels:	0
DB:	21	Gaps:	0

US-09-774-490-1 (1-2709) x AAB28379 (1-751)

OY 1028 AGAAGCTGTGTAATTAATGACAACTTCCTCAAGCTGCTG 1072

DB 274 ArgSerLeuValAsnLysTrpThrRhrPheLeuLysAlaArgLeu 288

RESULT 69

ID AAG62728 standard; peptide; 751 AA.

XX AAG62728;

XX 17-SEP-2001 (first entry)

XX Amino acid sequence of human semaphorin Sema3C.

XX Hanatoxin; taramula; toxin; voltage-gated potassium channel;
XX voltage-gated calcium channel; hanatoxin-like sequence; HTLS;
XX semaphorin; dorsal root ganglion repulsion; growth cone collapse.

XX Homo sapiens.

XX WO200138491-A2.

XX 31-MAY-2001.

XX 07-NOV-2000; 2000MO-US41943.

XX 08-NOV-1999; 99US-0164056.

XX (GENO) GEN HOSPITAL CORP.

XX Behar O, Woolf CJ;

XX WPI; 2001-451494/48.

XX Polypeptide sequences that encompass the hanatoxin-like sequences of
XX semaphorins, useful as a drugs to treat any condition or disease that
XX is characterized by abnormal calcium channel function -

XX Claim 6; Page 12; 29pp; English.

XX The present sequence represents a semaphorin. The specification
XX describes hanatoxin-like sequences (HTLS) found in the semaphorin
XX domain of mammalian secreted semaphorins. Hanatoxin is a taramula
XX toxin that selectively blocks some voltage-gated potassium and calcium
XX channels. The HTLS is responsible for the dorsal root ganglion repulsion
XX and growth cone collapse activities associated with semaphorins.
XX Polypeptides containing HTLS can be used to modulate the activity of
XX calcium channels. The peptides can also be used as an antigen to
XX generate antibodies that can then be used to modulate the activity of
XX calcium channels by inactivating naturally occurring channel ligands.
XX The peptides or antibodies can be used as drugs to treat any condition
XX or disease that is characterized by abnormal calcium channel function.

XX Sequence 751 AA;

Alignment Scores:

Pred. No.:	6,13e-06	Length:	751
Score:	15.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.68%	Indels:	0
DB:	22	Gaps:	0

US-09-774-490-1 (1-2709) x AAG62728 (1-751)

QY 1028 AGAAGTCTGTAATGAACAACATTCTCAAGCTGCTG 1072
 |||||
 DB 274 ArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeu 288

RESULT 70

ABP68623 standard; Protein; 751 AA.

AC ABP68623;

DT 14-JAN-2003 (first entry)

DE Human pancreatic cancer expressed protein SEQ ID NO 172.

XX Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;

XX Cytostatic; tumour.

XX Homo sapiens.

XX WO200260317-A2.

XX 08-AUG-2002.

XX 30-JAN-2002; 2002MO-US02781.

XX 30-JAN-2001; 2001US-265305P.

XX 31-JAN-2001; 2001US-265682P.

XX 09-FEB-2001; 2001US-267568P.

XX 21-MAR-2001; 2001US-278651P.

XX 28-APR-2001; 2001US-287112P.

XX 16-MAY-2001; 2001US-291631P.

XX 12-JUL-2001; 2001US-305484P.

XX 20-AUG-2001; 2001US-313999P.

XX 27-NOV-2001; 2001US-333626P.

XX (CORI-) CORIXA CORP.

XX Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;

XX WPI; 2002-627435/67.

XX N-PSDB; ABV94769.

XX New isolated polynucleotide and pancreatic tumor polypeptides, useful

XX for diagnosing, preventing and/or treating cancer, particularly

XX pancreatic cancer

XX Claim 2; SEQ ID NO 172; 300pp + Sequence Listing; English.

XX The invention relates to an isolated polynucleotide (I) comprising: (a)

XX any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145);

XX (b) complements of (a); (c) sequences consisting of at least 20

XX contiguous residues of (a); (d) sequences that hybridize to (a), under

XX moderately stringent conditions; (e) sequences having at least 75% or 90%

XX identity to (a); or (f) degenerate variants of (a). Polypeptides

XX (ABP68596-ABP68637) encoded by (I) and oligonucleotide can be used to

XX detect cancer in a patient and compositions comprising polypeptides,

XX polynucleotides, antibodies, fusion proteins, T cell populations and

XX antigen presenting cells expressing the polypeptide are useful in

XX treating pancreatic cancer and stimulating an immune response. The

XX polynucleotides can be used as probes or primers for nucleic acid

XX hybridisation, in the design and preparation of ribozyme molecules for

XX inhibiting expression of the tumour polypeptides and proteins in the

CC tumour cells, in vaccines and for gene therapy.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcc_sequences.

SQ Sequence 751 AA;

Alignment Scores:

Pred. No.:	6,13e-06	Length:	751
Score:	15.00	Matches:	15
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.68%	Indels:	0
DB:	23	Gaps:	0

US-09-774-490-1 (1-2709) x ABP68623 (1-751)

QY 1028 AGAAGTCTGTAATGAACAACATTCTCAAGCTGCTG 1072
 |||||
 DB 274 ArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeu 288

RESULT 71

AAV21315 standard; Protein; 14 AA.

AC AAV21315;

DT 22-JUL-1999 (first entry)

DE Human semaphorin III mutant protein fragment 48.

XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;

XX frameshift mutation; age-related disease; neurodegenerative disorder;

XX Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;

XX Huntington's disease; multiple sclerosis; alcoholic liver disease;

XX diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;

XX ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;

XX neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;

XX glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;

XX bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSF-A;

XX high mobility group protein-C; neuroendocrine specific protein A.

XX Synthetic.

XX Homo sapiens.

XX WO9845322-A2.

XX 15-OCT-1998.

XX 02-APR-1998; 98WO-1B00705.

XX 10-APR-1997; 97US-0043163.

XX (UYUT-) RIJXSUNIV UTRECHT.

XX (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.

XX (UYRO-) UNIV ROTTERDAM ERASMUS.

XX Burbach JPH, Grosveld FG, Van Leeuwen FW;

XX WPI; 1998-609901/51.

XX N-PSDB; AAX75767.

XX Diagnosing disease by detecting frameshift mutations in RNA or

XX corresponding protein mutations - used to diagnose cancer and

XX neurological diseases, particularly Alzheimer's disease, and also

XX for treatment and prevention with specific ribozymes or wild-type

XX RNA

XX Disclosure; Figure 16; 25pp; English.

XX This invention describes a novel method for the diagnosis of a disease

XX caused by, or associated with, an RNA molecule that has a frameshift

XX mutation. The method is used to diagnose age-related diseases, especially

CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's disease, Down's syndrome, myotonic dystrophy, Huntington's disease, CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II CC and many others listed) or susceptibility to these disorders. The method CC allows a definitive diagnosis of Alzheimer's disease in living patients, CC at an early stage. It is based on the observation that disease may be CC caused by mutations in RNA rather than DNA. The invention describes the CC use of neuronal system RNA molecules, specifically proteins including CC beta-amyloid precursor protein (beta-APP), the microtubule associated CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M, CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group CC protein-C (HMGP-C) and neuroendocrine specific protein A.

SO Sequence 14 AA;

Alignment Scores:

Pred. No.:	0.000108	Length:	14
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.56%	Indels:	0
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) x AAY21276 (1-14)

QY 946 AATGCAATGATGAGAACTCTGGAAAGCTACTACGC 987

Db 1 LysCysAsnArgTrpArgThrLeuTrpLysSerTySerArg 14

RESULT 72

AAY21276 standard; Protein; 14 AA.

AC AAY21276;

DT 22-JUL-1999 (first entry)

DE Human semaphorin III mutant protein fragment 9.

XX Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;
 KW frameshift mutation; age-related disease; neurodegenerative disorder;
 KW Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;
 KW Huntington's disease; multiple sclerosis; alcoholic liver disease;
 KW diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;
 KW ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;
 KW neurofilament-F; presentin I; presentin II; cellular tumour antigen;
 KW glial fibrillary acidic protein; GFAP; p53; semaphorin III; HUPF-1;
 KW bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; HMGP-C; NSP-A;
 KW high mobility group protein-C; neuroendocrine specific protein A.

XX Synthetic.

OS Homo sapiens.

XX MO9845322-A2.

XX 15-OCT-1998.

XX 02-APR-1998; 98MO-IB00705.

XX 10-APR-1997; 97US-0043163.

XX (UUT-) RIJKSUNIV UTRECHT.

XX (ROYA-) ROYAL NETHERLANDS ACAD ARTS & SCI.

XX (UVO-) UNIV ROTTERDAM ERASMUS.

XX Burbach JPH, Grosveld FG, Van Leeuwen FW;

XX MPI, 1998-609901/51.

XX DR N-ESDB; AAX75767.

PT Diagnosing disease by detecting frameshift mutations in RNA or
 PT corresponding protein mutations - used to diagnose cancer and
 PT neurological diseases, particularly Alzheimer's disease, and also
 PT for treatment and prevention with specific ribozymes or wild-type
 RNA

PS Disclosure; Figure 16; 258pp; English.

CC This invention describes a novel method for the diagnosis of a disease
 CC caused by, or associated with, an RNA molecule that has a frameshift
 CC mutation. The method is used to diagnose age-related diseases, especially
 CC cancer and a wide range of neurodegenerative disorders (e.g. Alzheimer's
 CC disease, Down's syndrome, myotonic dystrophy, Huntington's disease,
 CC multiple sclerosis, alcoholic liver disease, diabetes mellitus type II
 CC and many others listed) or susceptibility to these disorders. The method
 CC allows a definitive diagnosis of Alzheimer's disease in living patients,
 CC at an early stage. It is based on the observation that disease may be
 CC caused by mutations in RNA rather than DNA. The invention describes the
 CC use of neuronal system RNA molecules, specifically proteins including
 CC beta-amyloid precursor protein (beta-APP), the microtubule associated
 CC proteins Tau and Big Tau, ubiquitin B, apolipoprotein E, microtubule
 CC associated protein 2 (MAP2), neurofilament-L, neurofilament-M,
 CC neurofilament-F, presentin I, presentin II, glial fibrillary acidic
 CC protein (GFAP), the cellular tumour antigen p53, B-cell leukemia/lymphoma
 CC 2 (bcl-2) proto-oncogene, semaphorin III, HUPF-1, high mobility group
 CC protein-C (HMGP-C) and neuroendocrine specific protein A.

SO Sequence 14 AA;

Alignment Scores:

Pred. No.:	0.000108	Length:	14
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.56%	Indels:	0
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) x AAY21276 (1-14)

QY 996 GTCAGATATGCAAGATGACTTTGGAGGGCACAGAGTCTGG 1037

Db 1 ValArgTyrAlaArgMetThrLeuGlnGlyThrGluValTrp 14

RESULT 73

AAY27127 standard; Protein; 777 AA.

AC AAY27127;

DT 14-SEP-1999 (first entry)

DE Human brain tissue-derived polypeptide (clone OM007).

XX Brain tissue; human; bone marrow; umbilical cord venous endothelial cell;
 KW recombinant; diagnosis; treatment.

XX Homo sapiens.

XX MO9933873-A1.

XX 08-JUL-1999.

XX 25-DEC-1998; 98MO-JP05952.

XX 26-DEC-1997; 97JP-0358811.

XX (ONVO) ONO PHARM CO LTD.

XX Key Location/Qualifiers

XX Peptide 1..36

XX Protein 37..777

XX /note= "mature protein"

XX PI Fukushima D, Shibayama S, Tada H;
XX WPI: 1999-419088/35.
DR N-PSDB; AAX89112, AAX89113.
XX
PT New adult human brain tissue-produced polypeptides useful for
PT diagnosis and treatment
XX
PS Claim 1; Page 36-39; 86pp; Japanese.
XX
CC The invention provides polypeptides (AAV27127-Y27133) produced by human
CC adult brain tissue, human bone marrow or a human umbilical cord venous
CC endothelial cell. Host cells transformed with vectors comprising the
CC nucleic acids encoding the polypeptides are used for the recombinant
CC expression of the polypeptides. The polypeptides can be used in
CC diagnosis, treatment and basic studies, with wide applications in
CC treatment depending on the activity to be aimed at. Sequences
CC AAX89112-135 represent nucleic acids encoding the polypeptides.
XX
SQ Sequence 777 AA;

Alignment Scores:
Pred. No.: 6,42e-05 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 20 Gaps: 0

US-09-774-490-1 (1-2709) x AAY27127 (1-777)

QY 1040 AATAATGACACATTCCTCAAGCTCGTCGATTGGCTCA 1081
Db 298 Aenlystrpnrhrpheuylsalaaguenilecysser 311

RESULT 74
AAV99427
ID AAY99427 standard; Protein; 777 AA.
XX
AC AAY99427;
XX
DT 08-AUG-2000 (first entry)
XX
DE Human PRO1491 (UNQ760) amino acid sequence SEQ ID NO:310.
XX
KW Human; PRO polypeptide; membrane bound protein; receptor; diagnosis;
KW transmembrane; secretion; immunoadhesion; pharmaceutical; screening.
OS Homo sapiens.
XX
PN WO200012708-A2.
XX
PD 09-MAR-2000.
XX
PF 01-SEP-1999; 99WC-US20111.
XX
XX 01-SEP-1998; 98US-0098716.
PR 01-SEP-1998; 98US-0098749.
PR 01-SEP-1998; 98US-0098750.
PR 02-SEP-1998; 98US-0098803.
PR 02-SEP-1998; 98US-0098821.
PR 02-SEP-1998; 98US-0098843.
PR 09-SEP-1998; 98US-0099536.
PR 09-SEP-1998; 98US-0099596.
PR 09-SEP-1998; 98US-0099598.
PR 09-SEP-1998; 98US-0099602.
PR 09-SEP-1998; 98US-0099642.
PR 10-SEP-1998; 98US-0099741.
PR 10-SEP-1998; 98US-0099754.
PR 10-SEP-1998; 98US-0099763.
PR 10-SEP-1998; 98US-0099792.
PR 10-SEP-1998; 98US-0099808.

PR 10-SEP-1998; 98US-0099812.
PR 10-SEP-1998; 98US-0099815.
PR 10-SEP-1998; 98US-0099816.
PR 15-SEP-1998; 98US-0100385.
PR 15-SEP-1998; 98US-0100388.
PR 15-SEP-1998; 98US-0100390.
PR 16-SEP-1998; 98US-0100584.
PR 16-SEP-1998; 98US-0100627.
PR 16-SEP-1998; 98US-0100661.
PR 16-SEP-1998; 98US-0100662.
PR 16-SEP-1998; 98US-0100664.
PR 17-SEP-1998; 98US-0100683.
PR 17-SEP-1998; 98US-0100684.
PR 17-SEP-1998; 98US-0100710.
PR 17-SEP-1998; 98US-0100711.
PR 17-SEP-1998; 98US-0100919.
PR 17-SEP-1998; 98US-0100930.
PR 18-SEP-1998; 98US-0100848.
PR 18-SEP-1998; 98US-0100849.
PR 18-SEP-1998; 98US-0101014.
PR 18-SEP-1998; 98US-0101068.
PR 18-SEP-1998; 98US-0101071.
PR 22-SEP-1998; 98US-0101279.
PR 23-SEP-1998; 98US-0101471.
PR 23-SEP-1998; 98US-0101472.
PR 23-SEP-1998; 98US-0101474.
PR 23-SEP-1998; 98US-0101475.
PR 23-SEP-1998; 98US-0101476.
PR 23-SEP-1998; 98US-0101477.
PR 23-SEP-1998; 98US-0101479.
PR 24-SEP-1998; 98US-0101738.
PR 24-SEP-1998; 98US-0101741.
PR 24-SEP-1998; 98US-0101743.
PR 24-SEP-1998; 98US-0101915.
PR 24-SEP-1998; 98US-0101916.
PR 29-SEP-1998; 98US-0102207.
PR 29-SEP-1998; 98US-0102240.
PR 29-SEP-1998; 98US-0102307.
PR 29-SEP-1998; 98US-0102330.
PR 29-SEP-1998; 98US-0102331.
PR 30-SEP-1998; 98US-0102484.
PR 30-SEP-1998; 98US-0102487.
PR 30-SEP-1998; 98US-0102570.
PR 30-SEP-1998; 98US-0102571.
PR 01-OCT-1998; 98US-0102684.
PR 01-OCT-1998; 98US-0102687.
PR 02-OCT-1998; 98US-0102965.
PR 06-OCT-1998; 98US-0103258.
PR 06-OCT-1998; 98US-0103449.
PR 07-OCT-1998; 98US-0103314.
PR 07-OCT-1998; 98US-0103315.
PR 07-OCT-1998; 98US-0103328.
PR 07-OCT-1998; 98US-0103328.
PR 07-OCT-1998; 98US-0103395.
PR 07-OCT-1998; 98US-0103396.
PR 07-OCT-1998; 98US-0103401.
PR 08-OCT-1998; 98US-0103633.
PR 08-OCT-1998; 98US-0103678.
PR 08-OCT-1998; 98US-0103679.
PR 08-OCT-1998; 98US-0103711.
PR 14-OCT-1998; 98US-0104257.
PR 20-OCT-1998; 98US-0104987.
PR 20-OCT-1998; 98US-0105000.
PR 20-OCT-1998; 98US-0105002.
PR 21-OCT-1998; 98US-0105104.
PR 22-OCT-1998; 98US-0105169.
PR 22-OCT-1998; 98US-0105266.
PR 26-OCT-1998; 98US-0105693.
PR 26-OCT-1998; 98US-0105694.
PR 27-OCT-1998; 98US-0105807.
PR 27-OCT-1998; 98US-0105881.
PR 27-OCT-1998; 98US-0105882.
PR 27-OCT-1998; 98US-0106062.
PR 28-OCT-1998; 98US-0106023.

PR 28-OCT-1998; 98US-0106029.
 PR 28-OCT-1998; 98US-0106030.
 PR 28-OCT-1998; 98US-0106032.
 PR 28-OCT-1998; 98US-0106033.
 PR 28-OCT-1998; 98US-0106178.
 PR 29-OCT-1998; 98US-0106248.
 PR 29-OCT-1998; 98US-0106384.
 PR 29-OCT-1998; 98US-0106500.
 PR 30-OCT-1998; 98US-0106464.
 PR 30-OCT-1998; 98US-0106566.
 PR 03-NOV-1998; 98US-0106802.
 PR 03-NOV-1998; 98US-0106905.
 PR 03-NOV-1998; 98US-0106919.
 PR 03-NOV-1998; 98US-0106932.
 PR 03-NOV-1998; 98US-0106934.
 PR 10-NOV-1998; 98US-0107783.
 PR 17-NOV-1998; 98US-0108775.
 PR 17-NOV-1998; 98US-0108779.
 PR 17-NOV-1998; 98US-0108787.
 PR 17-NOV-1998; 98US-0108788.
 PR 17-NOV-1998; 98US-0108801.
 PR 17-NOV-1998; 98US-0108802.
 PR 17-NOV-1998; 98US-0108806.
 PR 17-NOV-1998; 98US-0108807.
 PR 17-NOV-1998; 98US-0108867.
 PR 18-NOV-1998; 98US-0108925.
 PR 18-NOV-1998; 98US-0108848.
 PR 18-NOV-1998; 98US-0108849.
 PR 18-NOV-1998; 98US-0108850.
 PR 18-NOV-1998; 98US-0108851.
 PR 18-NOV-1998; 98US-0108852.
 PR 18-NOV-1998; 98US-0108858.
 PR 18-NOV-1998; 98US-0108904.
 XX
 PA (GENTH) GENENTECH INC.
 PI Baker K, Goddard A, Gurney AL, Smith V, Watanabe CK, Wood WI;
 XX WPI: 2000-237871/20.
 DR N-PSDB; AAA37109.
 DR
 PT New mammalian DNA sequences encoding transmembrane, receptor or
 PT secreted PRO polypeptides, useful for screening of potential peptide or
 PT small molecule inhibitors of the relevant receptor/ligand interactions
 XX
 PS Claim 12; Fig 176; 773pp; English.
 CC AAA37022 to AAA37144 encode the new isolated human transmembrane,
 CC receptor or secreted PRO polypeptides given in AAY9340 to AAY93462. The
 CC transmembrane and receptor PRO proteins can be used for screening of
 CC potential peptide or small molecule inhibitors of the relevant
 CC receptor/ligand interactions. The polypeptides and nucleotide sequences
 CC encoding then have various industrial applications, including uses as
 CC pharmaceutical and diagnostic agents. AAA37145 to AAA37330 represent
 CC PCR primers and hybridisation probes used in the isolation of the PRO
 CC polypeptides from the present invention.
 XX
 SQ Sequence 777 AA:

Alignment Scores: 6.42e-05 Length: 777
 Pred. No.: 14.00 Matches: 14
 Score: 100.00% Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 1.56% Gaps: 0
 DB: 21

US-09-774-490-1 (1-2709) x AAY99427 (1-777)

OY 1040 AATAATGGAACAATTCCTCAAGCTGCTGATTGCTCA 1081
 DB 238 AenlysrtptrhrthphenelysAlaArgleuileCySer 311

RESULT 75
 AAU29197
 ID AAU29197 standard; Protein; 777 AA.
 XX
 AC AAU29197;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human PRO polypeptide sequence #174.
 XX
 KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
 KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
 KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
 KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200168848-A2.
 XX
 PD 20-SEP-2001.
 XX
 PP 28-FEB-2001; 2001WO-US06520.
 XX
 PR 01-MAR-2000; 2000WO-US05601.
 PR 02-MAR-2000; 2000WO-US05841.
 PR 03-MAR-2000; 2000US-187202P.
 PR 06-MAR-2000; 2000US-186968P.
 PR 14-MAR-2000; 2000US-189320P.
 PR 14-MAR-2000; 2000US-189328P.
 PR 15-MAR-2000; 2000WO-US06884.
 PR 21-MAR-2000; 2000US-190828P.
 PR 21-MAR-2000; 2000US-191007P.
 PR 21-MAR-2000; 2000US-191048P.
 PR 21-MAR-2000; 2000US-191314P.
 PR 28-MAR-2000; 2000US-192655P.
 PR 29-MAR-2000; 2000US-193032P.
 PR 29-MAR-2000; 2000US-193053P.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 04-APR-2000; 2000US-194449P.
 PR 04-APR-2000; 2000US-194647P.
 PR 11-APR-2000; 2000US-195975P.
 PR 11-APR-2000; 2000US-196000P.
 PR 11-APR-2000; 2000US-196187P.
 PR 11-APR-2000; 2000US-196690P.
 PR 11-APR-2000; 2000US-196820P.
 PR 18-APR-2000; 2000US-198121P.
 PR 18-APR-2000; 2000US-198585P.
 PR 25-APR-2000; 2000US-199397P.
 PR 25-APR-2000; 2000US-199550P.
 PR 25-APR-2000; 2000US-199654P.
 PR 03-MAY-2000; 2000US-201516P.
 PR 17-MAY-2000; 2000WO-US13705.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 30-MAY-2000; 2000WO-US14941.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 05-JUN-2000; 2000US-209832P.
 PR 28-JUL-2000; 2000WO-US20710.
 PR 22-AUG-2000; 2000US-0644848.
 PR 24-AUG-2000; 2000WO-US23328.
 PR 08-NOV-2000; 2000WO-US30952.
 PR 01-DEC-2000; 2000WO-US34678.
 PR 20-DEC-2000; 2000WO-US34956.
 XX
 PA (GENTH) GENENTECH INC.
 PI Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
 PI Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
 XX WPI: 2001-602746/68.
 DR N-PSDB; AAS46098.
 DR
 PT Novel nucleic acids encoding PRO polypeptides, used to diagnose the
 PT presence of tumours, such as prostate and breast tumours, in mammals and

PT to screen for modulators of the compounds -

XX
PS Claim 11, Fig 348; 774pp; English.
XX

CC Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
CC The PRO polypeptides and their associated nucleic acids can be used to
CC detect the presence of a tumour in a mammal by comparing the level of
CC expression of a PRO polypeptide in a test sample of cells from the animal
CC and a control sample of normal cells, whereby a higher level of
CC expression in the test sample indicates the presence of a tumour in the
CC mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
CC and rabbits but are preferably human. The polypeptides can be used to
CC stimulate tumour necrosis factor (TNF) alpha release from human blood,
CC when contacted with it. A specific polypeptide can be used to stimulate
CC the proliferation or differentiation of chondrocyte cells. The PRO
CC proteins can be used to determine the presence of tumours and also
CC susceptibility to tumour development, particularly adrenal, lung, colon,
CC breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC can be used for genetic analysis of individuals with genetic disorders.
XX

SQ Sequence 777 AA:

Alignment Scores:
Pred. No.: 6.42e-05 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 22 Gaps: 0

US-09-774-490-1 (1-2709) x AAU29197 (1-777)

QY 1040 AATPAATGGACACATTCTCAAGCTCGTGTGATTGCTCA 1081
|||
Db 298 AsnLysIrpIhrThrPheLeuLysAlaArgLeuIleCysSer 311

Search completed: August 3, 2003, 10:49:42
Job time : 116.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comphen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: August 3, 2003, 10:47:09 ; Search time 57 Seconds
(without alignments)
11288.429 Million cell updates/sec

Title: US-09-774-490-1
Perfect score: 895
Sequence: 1 aatctttatctatcagatg.....agcctttttctctaacc 2709

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 451899 seqs, 118759770 residues

Word size: 1

Total number of hits satisfying chosen parameters: 847952

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODEL=frame+.n2p.model -DEV=xlh
-O=/cgn2_1/USFTO_spool/US09774490/runac_03082003_102929_4564/app_query.fasta.1.2887
-DB=Published Applications_AA -QMT=fastan -SUFFIX=oligo.rapb -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo
-TRANS=human40.cdi -LIST=100 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=75 -MODE=LOCAL -OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USFR=US09774490 @CGN 1.1.85 @runac_03082003_102929_4564
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARM TIMEOUT=30 -THREDS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published Applications_AA:*

1: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubppaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubppaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubppaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubppaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB.pep.*
17: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
------------	-------	--------------------	----	-------------

71	86.1	771	15	US-10-097-340-284	Sequence 284, App
72	86.1	771	15	US-10-262-538-10	Sequence 10, App1
73	86.1	771	15	US-09-864-761-47112	Sequence 47112, A
	5.9	53	9	US-09-864-761-44075	Sequence 44075, A
	4.6	41	9	US-09-864-761-44552	Sequence 44552, A
	3.7	33	9	US-09-864-761-35708	Sequence 35708, A
	1.9	7	9	US-09-731-179-2	Sequence 2, App11
	1.9	17	9	US-09-813-290-4	Sequence 4, App11
	1.9	17	9	US-09-813-290-2	Sequence 12, App11
	1.7	15	15	US-10-262-538-12	Sequence 12, App1
	1.7	15	14	US-10-114-833-214	Sequence 214, App
	1.7	15	15	US-10-060-036-172	Sequence 172, App
	1.7	15	15	US-10-205-823-359	Sequence 359, App
	1.7	15	15	US-10-262-538-14	Sequence 14, App1
	1.6	14	15	US-09-946-374-110	Sequence 310, App
	1.6	14	14	US-10-052-586-348	Sequence 348, App
	1.6	14	15	US-10-174-580-348	Sequence 348, App
	1.6	14	15	US-10-176-758-348	Sequence 348, App
	1.6	14	15	US-10-175-737-348	Sequence 348, App
	1.6	14	15	US-10-173-706-348	Sequence 348, App
	1.6	14	15	US-10-175-738-348	Sequence 348, App
	1.6	14	15	US-10-175-732-348	Sequence 348, App
	1.6	14	15	US-10-176-482-348	Sequence 348, App
	1.6	14	15	US-10-176-757-348	Sequence 348, App
	1.6	14	15	US-10-176-913-348	Sequence 348, App
	1.6	14	15	US-10-176-913-348	Sequence 348, App
	1.6	14	15	US-10-180-552-348	Sequence 348, App
	1.6	14	15	US-10-180-557-348	Sequence 348, App
	1.6	14	15	US-10-173-700-348	Sequence 348, App
	1.6	14	15	US-10-174-572-348	Sequence 348, App
	1.6	14	15	US-10-174-579-348	Sequence 348, App
	1.6	14	15	US-10-174-582-348	Sequence 348, App
	1.6	14	15	US-10-174-588-348	Sequence 348, App
	1.6	14	15	US-10-175-729-348	Sequence 348, App
	1.6	14	15	US-10-175-740-348	Sequence 348, App
	1.6	14	15	US-10-175-743-348	Sequence 348, App
	1.6	14	15	US-10-176-488-348	Sequence 348, App
	1.6	14	15	US-10-176-492-348	Sequence 348, App
	1.6	14	15	US-10-176-747-348	Sequence 348, App
	1.6	14	15	US-10-176-750-348	Sequence 348, App
	1.6	14	15	US-10-176-985-348	Sequence 348, App
	1.6	14	15	US-10-176-987-348	Sequence 348, App
	1.6	14	15	US-10-176-992-348	Sequence 348, App
	1.6	14	15	US-10-176-993-348	Sequence 348, App
	1.6	14	15	US-10-184-658-348	Sequence 348, App
	1.6	14	15	US-10-176-981-348	Sequence 348, App
	1.6	14	15	US-10-173-685-348	Sequence 348, App
	1.6	14	15	US-10-173-697-348	Sequence 348, App
	1.6	14	15	US-10-173-705-348	Sequence 348, App
	1.6	14	15	US-10-174-576-348	Sequence 348, App
	1.6	14	15	US-10-174-585-348	Sequence 348, App
	1.6	14	15	US-10-174-586-348	Sequence 348, App
	1.6	14	15	US-10-175-747-348	Sequence 348, App
	1.6	14	15	US-10-176-481-348	Sequence 348, App
	1.6	14	15	US-10-176-485-348	Sequence 348, App
	1.6	14	15	US-10-176-487-348	Sequence 348, App
	1.6	14	15	US-10-176-493-348	Sequence 348, App
	1.6	14	15	US-10-176-493-348	Sequence 348, App
	1.6	14	15	US-10-176-756-348	Sequence 348, App
	1.6	14	15	US-10-176-911-348	Sequence 348, App
	1.6	14	15	US-10-176-919-348	Sequence 348, App
	1.6	14	15	US-10-180-546-348	Sequence 348, App
	1.6	14	15	US-10-180-547-348	Sequence 348, App
	1.6	14	15	US-10-180-549-348	Sequence 348, App
	1.6	14	15	US-10-180-555-348	Sequence 348, App
	1.6	14	15	US-10-180-559-348	Sequence 348, App
	1.6	14	15	US-10-181-000-348	Sequence 348, App
	1.6	14	15	US-10-183-010-348	Sequence 348, App
	1.6	14	15	US-10-183-012-348	Sequence 348, App
	1.6	14	15	US-10-184-614-348	Sequence 348, App

```

74 14 1.6 777 15 US-10-184-623-348 Sequence 348, App
75 14 1.6 777 15 US-10-184-635-348 Sequence 348, App
76 14 1.6 777 15 US-10-184-637-348 Sequence 348, App
77 14 1.6 777 15 US-10-184-646-348 Sequence 348, App
78 14 1.6 777 15 US-10-184-647-348 Sequence 348, App
79 14 1.6 777 15 US-10-184-652-348 Sequence 348, App
80 14 1.6 777 15 US-10-187-594-348 Sequence 348, App
81 14 1.6 777 15 US-10-187-596-348 Sequence 348, App
82 14 1.6 777 15 US-10-187-745-348 Sequence 348, App
83 14 1.6 777 15 US-10-187-885-348 Sequence 348, App
84 14 1.6 777 15 US-10-187-886-348 Sequence 348, App
85 14 1.6 777 15 US-10-199-464-348 Sequence 348, App
86 14 1.6 777 15 US-10-196-756-348 Sequence 348, App
87 14 1.6 777 15 US-10-176-751-348 Sequence 348, App
88 14 1.6 777 15 US-10-176-760-348 Sequence 348, App
89 14 1.6 777 15 US-10-176-990-348 Sequence 348, App
90 14 1.6 777 15 US-10-180-541-348 Sequence 348, App
91 14 1.6 777 15 US-10-180-542-348 Sequence 348, App
92 14 1.6 777 15 US-10-180-548-348 Sequence 348, App
93 14 1.6 777 15 US-10-180-551-348 Sequence 348, App
94 14 1.6 777 15 US-10-180-998-348 Sequence 348, App
95 14 1.6 777 15 US-10-180-999-348 Sequence 348, App
96 14 1.6 777 15 US-10-183-013-348 Sequence 348, App
97 14 1.6 777 15 US-10-184-612-348 Sequence 348, App
98 14 1.6 777 15 US-10-184-616-348 Sequence 348, App
99 14 1.6 777 15 US-10-184-617-348 Sequence 348, App
100 14 1.6 777 15 US-10-184-622-348 Sequence 348, App

```

ALIGNMENTS

RESULT 1

```

US-10-097-340-284
Sequence 284, Application US/10097340
Publication No. US20030087250A1

```

GENERAL INFORMATION:

```

APPLICANT: John MONAHAN
APPLICANT: Manjula GANNAVAPU
APPLICANT: Sebastian HOERSCHE
APPLICANT: Shubhangi KAMATKAR
APPLICANT: Steve G. KOVATIS
APPLICANT: Rachel E. MEYERS
APPLICANT: Michael MORRISSEY
APPLICANT: Peter OLANDT
APPLICANT: Ami SEN
APPLICANT: Peter VERIBY
APPLICANT: Gordon B. MILLS
APPLICANT: Robert C. BAST, Jr.
APPLICANT: Karen LU
APPLICANT: Rosemarie SCHMANDT
APPLICANT: Xumei ZHAO
APPLICANT: Karen GLATT
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
FILE REFERENCE: MRI-030
CURRENT APPLICATION NUMBER: US/10/097,340
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/324,967
PRIOR FILING DATE: 2001/09/26
PRIOR APPLICATION NUMBER: 60/311,732
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/325,102
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/323,580
PRIOR FILING DATE: 2001-09-19
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FastSeq for Windows Version 4.0

```

```

; SEQ ID NO 284
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-284

Alignment Scores:
Pred. No.: 0
Score: 771.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 86.15%
DB: 15
Gaps: 0

US-09-774-490-1 (1-2709) x US-10-097-340-284 (1-771)

QY 200 ATGGGCTGTTAACTGAGATGCTGCTCTTTCTGGGAGATATTAATTACAGCAAGACA 259
DB 1 MetGLYTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuThrAlaArgAla 20

QY 260 AACATTCAGAAATGGGAGAAACAATGCGCAAGCTGGAATTTCTTACAAAGAAATGTTG 319
DB 21 AsnTYrGlnAsnGlyLVysAsnAsnValProaGleuLVysLeuSerTYrLVsGlnLeu 40

QY 320 GAATCCAAACAATGTATCATCTTTCATAGGCTTGCCCAAGCTCCAGTTATCATACCTTC 379
DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTYrTrHisThrPhe 60

QY 380 CTTTGGATGAGAAACGAGTAGGCTGTATGTTGAGCAAAAGATCATATTTTCATTC 439
DB 61 LeuLeuAspGluGluArgSerArgLeuTYrValGlyAlaLVysAspHisIlePheSerPhe 80

QY 440 GACCTGCTTAATATTAAGATTTTCAAAAGATTTGTCGCCAGTATCTTACACGAGAGA 499
DB 81 AspleuValAsnIleLVysAspPheGlnLVysIleValITrpProValSerTYrTrArgArg 100

QY 500 GATGAATGCAATGGGCTGGAAGAAACATCCGTAAGAATGCTTAATTTCAAGATA 559
DB 101 AspGluCysLVysTrpAlaGlyLVysAspIleLeuLVsGluCysAlaAsnPheIleLVysVal 120

QY 560 CTTAAGGCAATATATCAAGCTCATCTTGACCTGTGAAACGGGAGCTTTTCATCAATT 619
DB 121 LeuLVsAlaLVysAsnGlnThrHisLeuTYrAlaLVysGlyThrGlyAlaPheHisProIle 140

QY 620 TGCACCTCATTTGAATTTGACATCATCTCGAGCAATATTTTAAGTGGAGAACTCA 679
DB 141 CysTrpTYrIleGlnIleGlyHisIshisProGluAspAsnIlePheLVysLeuLVsSer 160

QY 680 CATTTTGAAGAGCGCGCTGGAGAGATCATATGACCTTAAGCTGTCAGACATCCCTT 739
DB 161 HisPheGluAsnGlyLVysGlyLVysSerProTYrAspProLVysLeuThrAlaSerLeu 180

QY 740 TTAATGATGAGAAATTAATTAATCTTGAACCTGAGCTGATTTTATGGGCGAGACTTTGCT 799
DB 181 LeuIleAspGlyGluLeuLVysSerGlyThrAlaAlaAspPheMetGlyLVysAspPheAla 200

QY 800 ATCTCCGAGCTCTGGGAGACACACACCAATCAGAGACAGACATGATCCAGGTGG 859
DB 201 IlePheArgThrLeuGlyHisIshisProIleArgGlnGlnHisAspSerArgTrp 220

QY 860 CTCATGATCCAAAGTTATGATGATGATCCACCTCATCTGAGAGATGACATCTCAAGAT 919
DB 221 LeuAsnAspProLVysPheIleSerHisIshisLeuIleSerGluSerAspAsnProGluAsp 240

QY 920 GACAAAGTAACTTTTCTTCGTTGAAATGCAATGATGAGAAACACTCTGGAAGCT 979
DB 241 AspLVysValTYrPhePhePheArgGluAsnAlaIleAspGlyGlnHisSerGlyLVysAla 260

QY 980 ACTCAGCTGAAGATGTCAGATATGCAAGATATGCTTTGAGGGCAGACAGAGTCTGGTG 1039
DB 261 ThrHisAlaArgIleGlyGlnIleCysLVysAsnAspPheGlyGlyHisArgSerLeuVal 280

QY 1040 AATAATGAGCAACAATCTCTCAAGCTGCTGATTTGCTCAGTCCAGGTCCAATGAC 1099

```

Db	281	AsnLysTrpThrThrPheLeuLysLeuAlaArgLeuIleCysSerValProGlyProAsnGly	300
Oy	1100	ATTGACACCTCATTTTGTATGAACTGCGAGAGTGTATTTCTTAAGAATCTTTAAAGATCCCTAAA	1159
Db	301	ILesApTrHnI.sPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys	320
Oy	1160	AATCCAGTGTGATATGAGAGTGTTTACGACTCCAGTAAACATTTTCAAGGGATCAGCGGTG	1219
Db	321	AsnProValValLysGlyValPheThrThrsSerLeuAsnIlePheLysGlySerLeuVal	340
Oy	1220	TGTATGTATACATGATGATGATGTGAAGAGGGGTGTTCTTGCTCCATATGCCCACAGGAT	1279
Db	341	CysMetLysSerMetSerAspValAlaArgValPheLeuGlyProGlyAlaIleAlaGAsp	360
Oy	1280	GGACCCAAACATTCATGGGTGGCTTATTCAGAAAGAGTCCCTTATCCACGGCCGAGAACT	1339
Db	361	GlyProAsnLysGlnTrpValProIleGlnLysArgValProLysProAspProGlyThr	380
Oy	1340	TGTCCAGCAAAACATTTGGTGTGTGTGTGACTCTCAAGAGGACCTTCCGTGATGATTATA	1399
Db	381	CysProSerLysTrpPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle	400
Oy	1400	ACCTTTGCAGAAAGTCATCCAGCCATGTATACATCCAGTGTTCCTATGAACAATCCGCCA	1459
Db	401	ThrPheAlaArgSerHnI.sProAlaMetLysAsnProValPheProMetAsnAsnAlaArgPro	420
Oy	1460	ATTAGTGATCAAAACGGATGTAAATTATCAATTTACAAAAATTGTCGTAGACCCGAGTGAT	1519
Db	421	IleValIleLysTrpAspValAsnTrpGlnPheThrGlnIleValIleAspArgValAsp	440
Oy	1520	GCAGAAAGTGCACGATGATGTTATGTATTATCGAAGCAGATGTTGGACCGGTCTTAAA	1579
Db	441	AlaGlnAspGlyGlnTrpAspValMetSerPheIleGlyThrAspValGlyThrValLeuLys	460
Oy	1580	GTAGTTTCAATTCCTAAGAGGACCTTGTGATGATTTTGAAGAGGTTCTGCGAAGAAATG	1639
Db	461	ValValSerIleProLysGlnThrTrpLysAspLeuGlnLysValLeuLeuGlnGlnMet	480
Oy	1640	ACAGTTTTCGGGAAACCGACTGCTAATTTACAGCAATGAGCTTTCCACTTAAGCAGCAACAA	1699
Db	481	ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln	500
Oy	1700	CTATATATATGGTTCAACGGCTGGGGTTGCCAGCTCCCTTTTAAACCGGTGTGATATTAC	1759
Db	501	LeuLysTrpIleGlySerThrAlaGlyValAlaGlnLeuProLeuHnI.sArgCysAspIleLysTr	520
Oy	1760	GGGAAAGCGGTGTGCTGAGTGTGTGCTGCCCGGAGCCCTTATCTGTGCTGGGAGAGTCT	1819
Db	521	GlyLysAlaCysAlaGluCysCysLeuHnI.sArgAspProLysCysAlaTrpAspGlySer	540
Oy	1820	GCATGTCTCGCTATTTTCCACTGCAGAAAGAGCGACMAAGCAGCAAGATATTAAGAAAT	1879
Db	541	AlaCysSerLysArgLysPheProThrAlaLysArgLysThrArgGlnAspIleArgAsn	560
Oy	1880	GGAGACCCACTGACTCACTGTTCAGACTTAAACCATGATATATACCATGGCCACAGCCCT	1939
Db	561	GlyAspProLeuTrHnI.sCysSerAspLeuHnI.sAspAsnHnI.sGlyLysLeuSerPro	580
Oy	1940	GAAGAAGAAATCATCTATGATGTGTGAGAGAAATGTAGCAATTTTGGAAATGAGCCGAAG	1999
Db	581	GlnGlnArgIleIleLysGlyValGlnAsnSerSerThrPheLeuGluCysSerProLys	600
Oy	2000	TTCGAGAGAGCGCTGTCTATTGGCAATTCAGAGCGAAATGAAGCGAAAGAAAGAG	2059
Db	601	SerGlnArgAlaLeuValLysTrpGlnPheGlnArgAlaGlnGlnLysGlyGlnGlu	620
Oy	2060	ATCAGAGGATGATCATATGATCAGAGCAGATCAAGGCTTTCGTACGATGCTTCAA	2119
Db	621	IleArgValAspAspHisIleIleLysGlyThrAspGlnGlyLeuLeuLeuAsnGlySerLeuGln	640
Oy	2120	CAGAGAGATTCAAGCAATTACCTTGCCATSCGGTGAACATGGGTTTATCAAACTCTT	2179

Db	641	GlnlysaSapserGlyAsnlyrLeuCyshSAlaValGlnHlEglYpHeileGlnTrleu	660
Qy	2180	CTTAAGGTAAACCCCTGGAGTCACTTTCACACAGACGATTTGGAGAACTTCTTCAATAAGAT	223
Db	661	LeuylsValThrLeuGlnValIleAspThrGlnHlSLeuGlnGluLeuLeuHlSlySAsp	660
Qy	2240	GATGATGAGATGAGCTCTTAAGACCAAGAAGATGTCCAAATAGCATGACACCTAGCCGAAG	229
Db	681	AspaSpGlyAspGlySerlySerlyThrlySgluMetSerAsnSerMetThrProSerGlnlyS	700
Qy	2300	GTCTGGTACAGACACTTTCATGCAAGCTCATCAACACCCCAATCTCAACACGATGATGAG	235
Db	701	ValTrpTyArgAspPheMetGlnLeuIleAsnHlSProAsnLeuAsnThrMetAspGlu	720
Qy	2360	TTCTGTGAACAAGTTTGGAAAAAGGGACCGAAAAACAACCGCGCAAGGGCCAGACATACC	241
Db	721	PheCySglnGlnValTrlTrpYAsArgAspArglySglnArgGlnArgGlnArgProGlyHlStr	740
Qy	2420	CCAGGAAACAGTAACAATGGAAGCACTTACAGAAATAATAGAAAGGTAGAAAACAGAGG	247
Db	741	ProGlyAsnSerAsnlySerlyThrlyShlSLeuGlnGlnValAsnlySglYArgAsnArgArg	760
Qy	2480	ACCCACGAATTTGAGAGGGCCACCCAGAGTGTTC	2512
Db	761	ThrHlSglnPheGlnArgAlaProArgSerVal	771
RESULT 2			
US-10-262-538-10			
; Sequence 10, Application US/10262538			
; Publication NO. US20030113324A1			
; GENERAL INFORMATION:			
; APPLICATION: Aitalo et al			
; TITLE OF INVENTION: NEUROFILIN/VEGF-C/VEGFR-3 MATERIALS AND METHODS			
; FILE REFERENCE: 28967/37564			
; CURRENT APPLICATION NUMBER: US/10/262, 538			
; CURRENT FILING DATE: 2002-09-30			
; NUMBER OF SEQ ID NOS: 33			
; SOFTWARE: PatentIn version 3.0			
; SEQ ID NO 10			
; LENGTH: 771			
; TYPE: PRF			
; ORGANISM: Homo sapiens			
US-10-262-538-10			
Alignment Scores:			
Pred. No.: 0			
Score: 771.00			
Percent Similarity: 100.00%			
Best Local Similarity: 100.00%			
Query Match: 86.15%			
DB: 15			
Gaps: 0			
US-09-774-490-1 (1-2709) x US-10-262-538-10 (1-771)			
Qy	200	ATGGGCTGTTAACTAGAGATTTGTCGTCTTTCTGGGAGTATTACTTACAGCAAGACA	259
Db	1	MetGlyTrpLeuThrArgIleValGlySLeuPheTrpGlyValLeuLeuThrAlaArgAla	20
Qy	260	AACATACAGATGGGAGAGAACATGTGCGCAAGCTGAATTATCTTCAACAAGAAATGTTG	319
Db	21	AsnTyArgIleAsnGlyLysAsnAsnValProArgLeuLyLeuSerTyrllySgluMetLeu	40
Qy	320	GAATCCAAACATGTGATCACTTTCAATGGCTTGGCCAAAGCTCCAGTTATCATCACTTC	379
Db	41	GluSerIleAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyrllyStrPhe	60
Qy	380	CTTTTGATGAGGAACGAGAGTGGCTGTATGTGTGGACAAAGATCACATATTTGATTC	439
Db	61	LeuLeuAspGlnGlnArgSerArgLeuTyrlValGlyAlaLysAspHlStrIlePheSerPhe	80
Qy	440	GACCTGTTATATCAAGGATTTTCAAAAGATTGTGTGGCCAGTATCTTACACACAGAGA	499
Db	81	AspLeuValAsnIleLysAspPheGlnLySleValTrpProValSerTyrlThrArgArg	1000

QY 500 GATGAATGCAAGTGGGCTGGAAGAAAGACATCTGTGAAGAAATGTGCTAATTTTCATCAAGGTA 559
 |||||
 Db 101 AspGluCySerIshTrpAlaGlyIshAspIleLeuIshGluCysAlaAsnPhelIshVal 120
 QY 560 CTTAAGGCATTAATCAAGATCTGACTGTGACCTGTGGAAAGGGGGCTTTTCATCCAAAT 619
 |||||
 Db 121 LeuIshAlaIshYshAsnGlnThrIshIshLeuIshIshCysGlyIshGlyAlaPhelIshProIle 140
 QY 620 TGCACCTACATTAATAATGGACATCATCTGAGAGAAATATTTTAACTGAGAGAACTCA 679
 |||||
 Db 141 CysThrIsh 160
 QY 680 CATTTGAAAACCGCGCTGGGAAGAGTCCATATGACCTTAAGCTGTGACAGCATCTT 739
 |||||
 Db 161 HisPhelGluAsnGlyIshArgIshIshSerProIshIshProIshIshIshIshIshIshIshIsh 180
 QY 740 TTAAATGATGAGAAATTAATCTCTGAACTGACGTGATTTTAAAGGGCGAGACTTTGCT 799
 |||||
 Db 181 LeuIshAspGlyIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 200
 QY 800 ATCTCCGAACTCTTGGGACCAACCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 859
 |||||
 Db 201 IlePhelArgThrLeuGlyIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 220
 QY 860 CTCATGATCCAAAGTTTCAATTAAGTCCACCTCATCTCAGAGAGTGAACAATCCTGAAGAT 919
 |||||
 Db 221 LeuAsnAspProIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 240
 QY 920 GACAAAGTACTCTTTCTTCGCTGAATGCAATGATGAGAGAGAGAGAGAGAGAGAGAGAG 979
 |||||
 Db 241 AspIshValIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 260
 QY 980 ACTCAAGCTAGAAATAGTATGCAAGAAATGACTTTGAGAGAGAGAGAGAGAGAGAGAGAG 1039
 |||||
 Db 261 ThrIshAlaIshArgIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 280
 QY 1040 AATAATGAGCAACATCTCTCAAGCTCGCTGATTTGCTCAGTGCAGGTCCAAATGAG 1099
 |||||
 Db 281 AsnIsh 300
 QY 1100 ATTGACATCTGATTTGATGAATGAGAGATGATTTCTTAATGAACTTTAAAGATCTTAA 1159
 |||||
 Db 301 IleAspThrIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 320
 QY 1160 AATCCAGTGTATATGAGAGGTTTAAAGCTTCAGAGAAATGCTTCAAGGATAGAGCGG 1219
 |||||
 Db 321 AsnProValIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 340
 QY 1220 TGTATGTATAGCATGATGATGAGAGAGGTTGCTTGTGCTATGATGATGATGATGAT 1279
 |||||
 Db 341 CysMetIsh 360
 QY 1280 GGAACCAATCATGATGAGGCTTATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1339
 |||||
 Db 361 GlyProAsnIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 380
 QY 1340 TGTCCAGAGAAACATTTGGTGTGTTGACTCTCAAGAAAGAGAGAGAGAGAGAGAGAGAG 1399
 |||||
 Db 381 CysProSerIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 400
 QY 1400 ACCTTGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1459
 |||||
 Db 401 ThrPhelIsh 420
 QY 1460 ATATGATCAAAACGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1519
 |||||
 Db 421 IleValIsh 440
 QY 1520 GGAAGAGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1579
 |||||
 Db 441 AlaGluAspGlyIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 460

QY 1580 GTAGTTTCATTTCTTAAGAGAGACTTGATATGATTTAGAAAGAGTTCTGTGGAAGAAATG 1639
 |||||
 Db 461 ValIsh 480
 QY 1640 ACAGTTTTTGGGAGAGCCGATCTGATTTTCAAGAAATGAGAGCTTTCCACTTAAGAGAGAGAG 1699
 |||||
 Db 481 ThrIsh 500
 QY 1700 CTATATATGTTGATCAAGAGCTGGGTTGCCAGCTCCCTTTTCAACCGGATGATATTTAC 1759
 |||||
 Db 501 LeuIsh 520
 QY 1760 GGAAGAGAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1819
 |||||
 Db 521 GlyIsh 540
 QY 1820 GCATGTTCTGCTATTTTCCCTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1879
 |||||
 Db 541 AlaCysSerIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 560
 QY 1880 GGAAGAGAGAGCTGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1939
 |||||
 Db 561 GlyAspProIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 580
 QY 1940 GAAGAGAGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1999
 |||||
 Db 581 GluIsh 600
 QY 2000 TGGCAGAGAGCGCTGTGCTATTTGGCAATTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2059
 |||||
 Db 601 SerGlnIsh 620
 QY 2060 ATCAGAGTGAATGATCATATATCATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2119
 |||||
 Db 621 IleArgValIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 640
 QY 2120 CAGAGAGATTCAGAGAAATTAATCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2179
 |||||
 Db 641 GlnIsh 660
 QY 2180 CTTAAGGTAACTCTGAGAGAGTCAATGACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2239
 |||||
 Db 661 LeuIsh 680
 QY 2240 GATGATGAGAGATGCTCTCAAG 2299
 |||||
 Db 681 AspAspGlyIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 700
 QY 2300 GTCTGTGATCAGAGACTTATGAGAGCTCATCAACCACTTCAACAGAGATGATGATGATGAT 2359
 |||||
 Db 701 ValIsh 720
 QY 2360 TTCTGTGAACAATTTGGAAG 2419
 |||||
 Db 721 PheCysGluIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 740
 QY 2420 CCAGGAG 2479
 |||||
 Db 741 ProGlyAsnIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIshIsh 760
 QY 2480 ACCCAGAGATTTGAG 2512
 |||||
 Db 761 ThrIsh

RESULT 3

US-09-864-761-47112
 ; Sequence 47112, Application US/09864761
 ; Patent No. US20020048763A1

GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.
 ; APPLICANT: Rank, David R.
 ; APPLICANT: Hanzel, David K.
 ; APPLICANT: Chen, Wensheng

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

Db 41 LeuysValThrIleuGluValIleAspIrnGluHisLeuGluIleuLeuHisIleAsp 60

Qy 2240 GATGATGAGATGCGCTTTAAGACCAAGAATGTCATATGATGACACCTAGCCAGAG 2239

Db 61 AspaBgIlyAspGlySerIyThrIySgluMetSerAsnSerMetIrnProSerGlnIly 80

Qy 2300 GCTGCGTACAGAGACTTTCATGACGCTCATCAACCCCAATTCATCAACGATGATGAG 2359

Db 81 ValTrpIyTrArgAspPheMetGlnLeuIleAsnHisProAsnIleuSnIrnMetAspGlu 100

Qy 2360 TTCTGTGAACAAGTTTGGAAAAAGGAGCGAAAAACAAGTCGGCAAAAGCCAGGACATACC 2419

Db 101 PheCySgluGlnValTrpIyAsTrArgAspArgIySgluIrnArgIlnrnProGlyHisThr 120

Qy 2420 CCAGGAACAGTACCAATATGAAAGCACTTACAGAAATAAGAAAGGTAGAAAACAGAGG 2479

Db 121 ProGlyAsnSerAsnIyTrpIyHisIleuGlnIleuHisIySglIyTrArgAsnArgArg 140

Qy 2480 ACCACGAAATTTGAAGAGGCGACCCAGGAGTGTTC 2512

Db 141 ThrHisGluPheGluArgIleProArgSerVal 151

RESULT 4

US-09-864-761-44075

Sequence 44075, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aemica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00667

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00664

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00669

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00668

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00663

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00670

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 2000-09-21

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: US 09/774,203

PRIOR FILING DATE: 2001-01-29

```
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 44075
/ LENGTH: 57
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC004848.1
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
/ OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALU 5.00e-25
US-09-864-761-44075

Alignment Scores:
Pred. No.: 2,73e-43 Length: 57
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0
DB: Gaps: 0

US-09-774-490-1 (1-2709) x US-09-864-761-44075 (1-57)

QY 311 GAAATTTGGATGCAACATGATGATCTTGAATGCGTGGCCAAAGCTCCAGTTAT 370
DB 5 G1uEluEnG1uEsrhEnvAl1leThPhaSnG1yEuAlaSnSerSerTy 24

QY 371 CATACCTCTTTGGATGAGAACGAGAGCTGATGTTGAGACAAAGATCACATA 430
DB 25 H1eThPhelEuEnuSpG1uG1uAygSerHryLeuTyAlG1yAlaYsaPhStle 44

QY 431 TTTTCATTGCACCTGTTAATATCAAGATTTTCAAG 469
DB 45 PhSeSerPhSePhSeValAen1leTyAspPhEg1n1ySe 57

RESULT 5
US-09-864-761-44552
/ Sequence 44552, Application US/09864761
/ Patent No. US20020048763A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ FILE REFERENCE: Aecm1ca-X-1
/ CURRENT APPLICATION NUMBER: US/09/864,761
/ CURRENT FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
```

```
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 44552
/ LENGTH: 41
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC006322.2
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.46
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.54
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56
/ OTHER INFORMATION: EST HUMAN HIT: BF700789.1, EVALU 1.00e-17
/ OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALU 1.00e-18
US-09-864-761-44552

Alignment Scores:
Pred. No.: 2,29e-31 Length: 41
Score: 41.00 Matches: 41
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.58% Indels: 0
DB: Gaps: 0

US-09-774-490-1 (1-2709) x US-09-864-761-44552 (1-41)

QY 1937 CCTGAAGAGATCATCTTATGCTGTAGAGATAGTACACATTTTGGATGAGTCCG 1996
DB 1 ProG1uG1uAyg1le1leTyrc1yAlG1uEnSerSerThPhE1uG1yCysSerPro 20

QY 1997 AAGTGGAGAGCGCTGTCTATTGCAATTCGAGAGCCAAAGAGAGGAAGAAGA 2056
DB 21 1ySsErG1nAygAlaLeuValTyrrpG1nPhEg1nAygAygEnG1uG1yG1u 40

QY 2057 GAG 2059
DB 41 G1u 41

RESULT 6
US-09-864-761-39708
/ Sequence 39708, Application US/09864761
/ Patent No. US20020048763A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Hanzel, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ FILE REFERENCE: Aecm1ca-X-1
/ CURRENT APPLICATION NUMBER: US/09/864,761
/ CURRENT FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2000-08-03
```

```

; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 39708
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
; OTHER INFORMATION: MAP TO AC004848.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUATION 4.00e-15
US-09-864-761-39708

Alignment Scores:
Pred. No.: 2,03e-23 Length: 33
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 9 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-864-761-39708 (1-33)

QY 533 AAGAGTGTCTTAATTTCTACAGGTACTTAAGCATATATACAGCTCTGTGACCC 592
DB 1 LysGUCysAlaAsnPhelIeYsValIeuYsAlaTyrAsnGlnThrHisIeuTyrAla 20
QY 593 TGTGACGGGGGCTTTTCATTCACATTTTGACCTTACATT 631
DB 21 CysGlyInrGlyAlaPheHisProIleCysInrTyrIle 33

RESULT 7
US-09-731-179-2
; Sequence 2, Application US/09731179
; Patent No. US20010049432A1
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L..
; APPLICANT: Foley, Kevin P.
```

```

; TITLE OF INVENTION: HUMAN SEMAPHORIN ZSMF-16
; FILE REFERENCE: 99-90
; CURRENT APPLICATION NUMBER: US/09/731,179
; CURRENT FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: US 60/169,238
; PRIOR FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 779
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-731-179-2

Alignment Scores:
Pred. No.: 1.11e-07 Length: 779
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 9 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-731-179-2 (1-779)

QY 1766 GCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTTGAGATGT 1816
DB 525 AlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTTPAspGly 541

RESULT 8
US-09-813-290-4
; Sequence 4, Application US/09813290
; Patent No. US20020042504A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Hilbun, Erin
; APPLICANT: Wang, Xiaoming
; APPLICANT: Donoho, Gregory
; APPLICANT: Scottille, John
; TITLE OF INVENTION: No. US20020042504A1 Human Secreted Proteins and Polynucleotides
; FILE REFERENCE: LEX-0151-USA
; CURRENT APPLICATION NUMBER: US/09/813,290
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/190,638
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: US 60/191,188
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: US 60/193,639
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 782
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-813-290-4

Alignment Scores:
Pred. No.: 1.11e-07 Length: 782
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 9 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-813-290-4 (1-782)

QY 1766 GCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTTGAGATGT 1816
DB 528 AlaCysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTTPAspGly 544

RESULT 9
```

```
US-09-813-290-2
; Sequence 2, Application US/09813290
; Patent No. US20020042504A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Wilganowski, Nathaniel L.
; APPLICANT: Turner, C. Alexander Jr.
; APPLICANT: Hilbun, Erin
; APPLICANT: Wang, Xiaoming
; APPLICANT: Donoho, Gregory
; APPLICANT: Scoville, John
; TITLE OF INVENTION: No. US20020042504A1 Human Secreted Proteins and Polynucleotides
; FILE REFERENCE: Lex-0151-USA
; CURRENT APPLICATION NUMBER: US/09/813,290
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: US 60/190,638
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: US 60/191,188
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: US 60/193,639
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 875
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-813-290-2

Alignment Scores:
Pred. No.: 1,1e-07 Length: 875
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 9 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-813-290-2 (1-875)

QY 1766 GGGTGTGAGAGTGGCTGCGCCGAGACCCCTTACTGTCGATGGATGGT 1816
DB 621 AAlAcYsAlAgLUCyScYsEuAlAaYgApPrIoYrCYeAlAtPaspGly 637

RESULT 10
US-10-262-538-12
; Sequence 12, Application US/10262538
; Publication No. US20030113324A1
; GENERAL INFORMATION:
; APPLICANT: Altalo et al
; TITLE OF INVENTION: NEUROFILIN/VEGF-C/VEGFR-3 MATERIALS AND METHODS
; FILE REFERENCE: 28967/37564
; CURRENT APPLICATION NUMBER: US/10/262,538
; CURRENT FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 749
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-538-12

Alignment Scores:
Pred. No.: 1,08e-05 Length: 749
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-262-538-12 (1-749)

QY 1028 AGAGTGTGTAATGAATGACAAACATTCTCAAGCTCGTCTG 1072
DB 274 ArgSerLeuValAsnLySerThrThrPheLeuYsAlaArgLeu 288

RESULT 12
US-10-060-036-172
; Sequence 172, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yuxiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 172
; LENGTH: 751
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-036-172

Alignment Scores:
Pred. No.: 1,08e-05 Length: 751
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
DB: 14 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-114-893-214 (1-751)

QY 1028 AGAGTGTGTAATGAATGACAAACATTCTCAAGCTCGTCTG 1072
DB 274 ArgSerLeuValAsnLySerThrThrPheLeuYsAlaArgLeu 288

RESULT 11
US-10-114-893-214
; Sequence 214, Application US/10114893
; Publication No. US20020193567A1
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaValle, Edward R.
; APPLICANT: Collins-Racie, Lisa A.
; APPLICANT: Evans, Cheryl
; APPLICANT: Weberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Bowman, Michael R.
; APPLICANT: Spaulding, Vikki
; APPLICANT: Carlin-Duckett, McKenough
; APPLICANT: Kelleher, Kerry S.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6000-10A
; CURRENT APPLICATION NUMBER: US/10/114,893
; CURRENT FILING DATE: 2002-04-02
; EARLIER APPLICATION NUMBER: 09/413,232
; EARLIER FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 321
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 214
; LENGTH: 751
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-893-214

Alignment Scores:
Pred. No.: 1,08e-05 Length: 751
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
DB: 14 Gaps: 0
```



```
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-060-036-172 (1-751)

Qy 1028 AGAAGCTGTGATGAATGAACAACTTCCTCAAGCTGCTG 1072
Db 274 ArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeu 288

RESULT 13
US-10-205-823-359
; Sequence 359, Application US/10205823
; Publication No. US20030108963A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Endege, Wilson O.
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Gorbatcheva, Bella
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Monsey, Angela M.
; APPLICANT: Glatz, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Anderson, Dustin
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITILE OF INVENTION: THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044
; CURRENT APPLICATION NUMBER: US/10/205,823
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 359
; LENGTH: 751
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-205-823-359

Alignment Scores:
Pred. No.: 1,08e-05 Length: 751
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-205-823-359 (1-751)

Qy 1028 AGAAGCTGTGATGAATGAACAACTTCCTCAAGCTGCTG 1072
Db 274 ArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeu 288

RESULT 14
US-10-262-538-14
; Sequence 14, Application US/10262538
; Publication No. US20030113324A1
; GENERAL INFORMATION:
; APPLICANT: Alltalo et al
; TITLE OF INVENTION: NEUROFILIN/VEGF-C/VEGFR-3 MATERIALS AND METHODS
```

```
; FILE REFERENCE: 28967/37564
; CURRENT APPLICATION NUMBER: US/10/262,538
; CURRENT FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 751
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-538-14

Alignment Scores:
Pred. No.: 1.08e-05 Length: 751
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-262-538-14 (1-751)

Qy 1028 AGAAGCTGTGATGAATGAACAACTTCCTCAAGCTGCTG 1072
Db 274 ArgSerLeuValAsnLysTrpThrPheLeuLysAlaArgLeu 288

RESULT 15
US-09-946-374-310
; Sequence 310, Application US/09946374
; Publication No. US20030073129A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Pan, James
; APPLICANT: Paoletti, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; ACIDS
; FILE REFERENCE: P2830P1C1
; CURRENT APPLICATION NUMBER: US/09/946,374
; CURRENT FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: 60/098716
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098723
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098749
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098750
; PRIOR FILING DATE: 1998-09-01
; PRIOR APPLICATION NUMBER: 60/098803
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098821
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/098843
; PRIOR FILING DATE: 1998-09-02
; PRIOR APPLICATION NUMBER: 60/099536
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099596
```

[illegible]

```
/ PRIOR FILING DATE: 1998-10-22
/ PRIOR APPLICATION NUMBER: 60/105266
/ PRIOR FILING DATE: 1998-10-22
/ PRIOR APPLICATION NUMBER: 60/105693
/ PRIOR FILING DATE: 1998-10-26
/ PRIOR APPLICATION NUMBER: 60/105694
/ PRIOR FILING DATE: 1998-10-26
/ PRIOR APPLICATION NUMBER: 60/105807

Alignment Scores:
Pred. No.: 0 000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 11 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-946-374-310 (1-777)

Qy 1040 AATAATGACAACTTCTCAAGCTGCTGATTGCTCA 1081
Db 298 AenlySTpThrThrPheueulySaIaArgleuIIeCySer 311

RESULT 16
US-10-052-586-348
/ Sequence 348, Application US/10052586
/ Publication No. US20020127584A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ TITLE OF INVENTION: ACIDS ENCODING THE SAME
/ FILE REFERENCE: P3430R1C1
/ CURRENT APPLICATION NUMBER: US/10/052,586
/ PRIOR FILING DATE: 2002-01-15
/ PRIOR APPLICATION NUMBER: 60/059263
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/059266
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/062250
/ PRIOR FILING DATE: 1997-10-17
/ PRIOR APPLICATION NUMBER: 60/063120
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/063121
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/063486
/ PRIOR FILING DATE: 1997-10-21
/ PRIOR APPLICATION NUMBER: 60/063540
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063541
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063544
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063564
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063734
/ PRIOR FILING DATE: 1997-10-29
/ PRIOR APPLICATION NUMBER: 60/063870
/ PRIOR FILING DATE: 1997-10-31
/ PRIOR APPLICATION NUMBER: 60/064103
/ PRIOR FILING DATE: 1997-10-31
/ PRIOR APPLICATION NUMBER: 60/065111
/ PRIOR FILING DATE: 1997-11-13
/ PRIOR APPLICATION NUMBER: 60/066120

/ PRIOR FILING DATE: 1997-11-21
/ PRIOR APPLICATION NUMBER: 60/066466
/ PRIOR FILING DATE: 1997-11-24
/ PRIOR APPLICATION NUMBER: 60/066772
/ PRIOR FILING DATE: 1997-11-24
/ PRIOR APPLICATION NUMBER: 60/069335
/ PRIOR FILING DATE: 1997-12-11
/ PRIOR APPLICATION NUMBER: 60/069425
/ PRIOR FILING DATE: 1997-12-12
/ PRIOR APPLICATION NUMBER: 60/069870
/ PRIOR FILING DATE: 1997-12-17
/ PRIOR APPLICATION NUMBER: 60/068017
/ PRIOR FILING DATE: 1997-12-18
/ PRIOR APPLICATION NUMBER: 60/07450
/ PRIOR FILING DATE: 1998-03-10
/ PRIOR APPLICATION NUMBER: 60/077632
/ PRIOR FILING DATE: 1998-03-11
/ PRIOR APPLICATION NUMBER: 60/077649
/ PRIOR FILING DATE: 1998-03-11
/ PRIOR APPLICATION NUMBER: 60/078886
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/078939
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/079664
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: 60/079786
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: 60/080107
/ PRIOR FILING DATE: 1998-03-31
/ PRIOR APPLICATION NUMBER: 60/080194
/ PRIOR FILING DATE: 1998-03-31
/ PRIOR APPLICATION NUMBER: 60/080327
/ PRIOR FILING DATE: 1998-04-01
/ PRIOR APPLICATION NUMBER: 60/080333
/ PRIOR FILING DATE: 1998-04-01
/ PRIOR APPLICATION NUMBER: 60/081049
/ PRIOR FILING DATE: 1998-04-08
/ PRIOR APPLICATION NUMBER: 60/081070
/ PRIOR FILING DATE: 1998-04-08
/ PRIOR APPLICATION NUMBER: 60/081195
/ PRIOR FILING DATE: 1998-04-09
/ PRIOR APPLICATION NUMBER: 60/081838
/ PRIOR FILING DATE: 1998-04-15
/ PRIOR APPLICATION NUMBER: 60/082568
/ PRIOR FILING DATE: 1998-04-21
/ PRIOR APPLICATION NUMBER: 60/082569
/ PRIOR FILING DATE: 1998-04-21
/ PRIOR APPLICATION NUMBER: 60/082704
/ PRIOR FILING DATE: 1998-04-22
/ PRIOR APPLICATION NUMBER: 60/082797
/ PRIOR FILING DATE: 1998-04-22
/ PRIOR APPLICATION NUMBER: 60/083322
/ PRIOR FILING DATE: 1998-04-28
/ PRIOR APPLICATION NUMBER: 60/083495
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083496
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083499
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083559
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/084366
/ PRIOR FILING DATE: 1998-05-05
/ PRIOR APPLICATION NUMBER: 60/084414
/ PRIOR FILING DATE: 1998-05-06
/ PRIOR APPLICATION NUMBER: 60/084639
/ PRIOR FILING DATE: 1998-05-07
/ PRIOR APPLICATION NUMBER: 60/084640
/ PRIOR FILING DATE: 1998-05-07
/ PRIOR APPLICATION NUMBER: 60/084643
/ PRIOR FILING DATE: 1998-05-07
/ PRIOR APPLICATION NUMBER: 60/085573
/ PRIOR FILING DATE: 1998-05-15
```

1	PRIOR APPLICATION NUMBER: 60/085579
2	PRIOR FILING DATE: 1998-05-15
3	PRIOR APPLICATION NUMBER: 60/085580
4	PRIOR FILING DATE: 1998-05-15
5	PRIOR APPLICATION NUMBER: 60/085582
6	PRIOR FILING DATE: 1998-05-15
7	PRIOR APPLICATION NUMBER: 60/085700
8	PRIOR FILING DATE: 1998-05-15
9	PRIOR APPLICATION NUMBER: 60/086023
10	PRIOR FILING DATE: 1998-05-18
11	PRIOR APPLICATION NUMBER: 60/086392
12	PRIOR FILING DATE: 1998-05-22
13	PRIOR APPLICATION NUMBER: 60/086486
14	PRIOR FILING DATE: 1998-05-22
15	PRIOR APPLICATION NUMBER: 60/087098
16	PRIOR FILING DATE: 1998-05-28
17	PRIOR APPLICATION NUMBER: 60/087208
18	PRIOR FILING DATE: 1998-05-28
19	PRIOR APPLICATION NUMBER: 60/087609
20	PRIOR FILING DATE: 1998-06-02
21	PRIOR APPLICATION NUMBER: 60/087759
22	PRIOR FILING DATE: 1998-06-02
23	PRIOR APPLICATION NUMBER: 60/087827
24	PRIOR FILING DATE: 1998-06-03
25	PRIOR APPLICATION NUMBER: 60/088025
26	PRIOR FILING DATE: 1998-06-04
27	PRIOR APPLICATION NUMBER: 60/088028
28	PRIOR FILING DATE: 1998-06-04
29	PRIOR APPLICATION NUMBER: 60/088029
30	PRIOR FILING DATE: 1998-06-04
31	PRIOR APPLICATION NUMBER: 60/088033
32	PRIOR FILING DATE: 1998-06-04
33	PRIOR APPLICATION NUMBER: 60/088157
34	PRIOR FILING DATE: 1998-06-05
35	PRIOR APPLICATION NUMBER: 60/088202
36	PRIOR FILING DATE: 1998-06-05
37	PRIOR APPLICATION NUMBER: 60/088212
38	PRIOR FILING DATE: 1998-06-05
39	PRIOR APPLICATION NUMBER: 60/088217
40	PRIOR FILING DATE: 1998-06-05
41	PRIOR APPLICATION NUMBER: 60/088326
42	PRIOR FILING DATE: 1998-06-04
43	PRIOR APPLICATION NUMBER: 60/088655
44	PRIOR FILING DATE: 1998-06-09
45	PRIOR APPLICATION NUMBER: 60/088722
46	PRIOR FILING DATE: 1998-06-10
47	PRIOR APPLICATION NUMBER: 60/088738
48	PRIOR FILING DATE: 1998-06-10
49	PRIOR APPLICATION NUMBER: 60/088740
50	PRIOR FILING DATE: 1998-06-10
51	PRIOR APPLICATION NUMBER: 60/088811
52	PRIOR FILING DATE: 1998-06-10
53	PRIOR APPLICATION NUMBER: 60/088824
54	PRIOR FILING DATE: 1998-06-10
55	PRIOR APPLICATION NUMBER: 60/088825
56	PRIOR FILING DATE: 1998-06-10
57	PRIOR APPLICATION NUMBER: 60/088826
58	PRIOR FILING DATE: 1998-06-10
59	PRIOR APPLICATION NUMBER: 60/088861
60	PRIOR FILING DATE: 1998-06-11
61	PRIOR APPLICATION NUMBER: 60/088900
62	PRIOR FILING DATE: 1998-06-12
63	PRIOR APPLICATION NUMBER: 60/089105
64	PRIOR FILING DATE: 1998-06-12
65	PRIOR APPLICATION NUMBER: 60/089512
66	PRIOR FILING DATE: 1998-06-16
67	PRIOR APPLICATION NUMBER: 60/089514
68	PRIOR FILING DATE: 1998-06-16
69	PRIOR APPLICATION NUMBER: 60/089538

```

; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 66/089558
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089653
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089908

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 14 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-052-586-348 (1-777)

OY 1040 AATAATGACAACATTCCTCAAGCTCGTGAATTGCTCA 1081
|||||
Db 298 AsnlystrpThrThrPheLeuYsAlaArgLeuIleCysSer 311

RESULT 17
US-10-174-590-348
; Sequence 348, Application US/10174590
; Publication No. US20030008352A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Deenoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C42
; CURRENT APPLICATION NUMBER: US/10/174,590
; CURRENT FILING DATE: 2002-06-18
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-590-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-174-590-348 (1-777)

OY 1040 AATAATGACAACATTCCTCAAGCTCGTGAATTGCTCA 1081
|||||
Db 298 AsnlystrpThrThrPheLeuYsAlaArgLeuIleCysSer 311

RESULT 18
US-10-176-758-348
; Sequence 348, Application US/10176758
; Publication No. US20030008353A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Deenoyers, Luc
; APPLICANT: Goddard, Audrey

```

APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C104
CURRENT APPLICATION NUMBER: US/10/176,758
CURRENT FILING DATE: 2002-06-21
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-758-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-758-348 (1-777)

OY 1040 AATAAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
DB 298 AenlystrpnrThrPheleuylsAlaArgleuileCysser 311

RESULT 19
US-10-175-737-348
Sequence 348, Application US/10175737
Publication No. US20030013153A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C50
CURRENT APPLICATION NUMBER: US/10/175,737
CURRENT FILING DATE: 2002-06-19
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-175-737-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-175-737-348 (1-777)

OY 1040 AATAAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
DB 298 AenlystrpnrThrPheleuylsAlaArgleuileCysser 311

RESULT 20
US-10-173-706-348
Sequence 348, Application US/10173706
Publication No. US20030022293A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C7
CURRENT APPLICATION NUMBER: US/10/173,706
CURRENT FILING DATE: 2002-06-17
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-173-706-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-173-706-348 (1-777)

OY 1040 AATAAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
DB 298 AenlystrpnrThrPheleuylsAlaArgleuileCysser 311

RESULT 21
US-10-175-738-348
Sequence 348, Application US/10175738
Publication No. US20030022294A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C45
CURRENT APPLICATION NUMBER: US/10/175,738
CURRENT FILING DATE: 2002-06-19
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT

```
; ORGANISM: Homo Sapien
US-10-175-738-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-175-738-348 (1-777)
QY 1040 AATAAATGACAACTTCTCAAGCTCGTGTGATTGCTCA 1081
Db 298 AsnlystrpThrPheLeuYsAlaArgLeuIleCysSer 311

RESULT 22
US-10-175-752-348
; Sequence 348, Application US/10175752
; Publication No. US20030022295A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C60
; CURRENT FILING DATE: 2002-06-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-752-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-175-752-348 (1-777)
QY 1040 AATAAATGACAACTTCTCAAGCTCGTGTGATTGCTCA 1081
Db 298 AsnlystrpThrPheLeuYsAlaArgLeuIleCysSer 311

RESULT 23
US-10-176-482-348
; Sequence 348, Application US/10176482
; Publication No. US20030022286A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
```

```
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C70
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-482-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-482-348 (1-777)
QY 1040 AATAAATGACAACTTCTCAAGCTCGTGTGATTGCTCA 1081
Db 298 AsnlystrpThrPheLeuYsAlaArgLeuIleCysSer 311

RESULT 24
US-10-176-757-348
; Sequence 348, Application US/10176757
; Publication No. US20030022297A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C86
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-757-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-757-348 (1-777)
QY 1040 AATAAATGACAACTTCTCAAGCTCGTGTGATTGCTCA 1081
Db 298 AsnlystrpThrPheLeuYsAlaArgLeuIleCysSer 311
```

```
RESULT 25
US-10-176-913-348
; Sequence 348, Application US/10176913
; Publication No. US20030022298A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C66
; CURRENT APPLICATION NUMBER: US/10/176,913
; CURRENT FILING DATE: 2002-06-20
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-913-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-913-348 (1-777)
QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
Db 298 AenlystrpThrPheleuylsAlaArgleuileCysSer 311

RESULT 26
US-10-180-552-348
; Sequence 348, Application US/10180552
; Publication No. US20030022300A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C153
; CURRENT APPLICATION NUMBER: US/10/180,552
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-552-348

Alignment Scores:
```

Alignment Scores:

```
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-180-552-348 (1-777)
QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
Db 298 AenlystrpThrPheleuylsAlaArgleuileCysSer 311

RESULT 27
US-10-180-557-348
; Sequence 348, Application US/10180557
; Publication No. US20030022301A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C147
; CURRENT APPLICATION NUMBER: US/10/180,557
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See file Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-557-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-180-557-348 (1-777)
QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
Db 298 AenlystrpThrPheleuylsAlaArgleuileCysSer 311

RESULT 28
US-10-173-700-348
; Sequence 348, Application US/10173700
; Publication No. US20030027262A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
```

```

; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C14
; CURRENT APPLICATION NUMBER: US/10/173,700
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-700-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-173-700-348 (1-777)

QY 1040 AATTAATGACAACTTCTCAAGCTCGTCTGATTGCTCA 1081
DB 298 AaLysTrpThrThrPheLeuYsAlaArgLeuIleCySer 311

RESULT 29
US-10-174-572-348
; Sequence 348, Application US/10174572
; Publication No. US20030027265A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C40
; CURRENT APPLICATION NUMBER: US/10/174,572
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-572-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-174-572-348 (1-777)

QY 1040 AATTAATGACAACTTCTCAAGCTCGTCTGATTGCTCA 1081
DB 298 AaLysTrpThrThrPheLeuYsAlaArgLeuIleCySer 311

RESULT 30
US-10-174-579-348
; Sequence 348, Application US/10174579
; Publication No. US20030027264A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C31
; CURRENT APPLICATION NUMBER: US/10/174,579
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-579-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-174-579-348 (1-777)

QY 1040 AATTAATGACAACTTCTCAAGCTCGTCTGATTGCTCA 1081
DB 298 AaLysTrpThrThrPheLeuYsAlaArgLeuIleCySer 311

RESULT 31
US-10-174-582-348
; Sequence 348, Application US/10174582
; Publication No. US20030027265A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C36
; CURRENT APPLICATION NUMBER: US/10/174,582
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-582-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
```


Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0
US-09-774-490-1 (1-2709) x US-10-174-588-348 (1-777)
QY 1040 AATAATGACAACTTCTCAAGCTCGTGTGATTGCTCA 1081
|||||
298 AenlystirpThrPheleuysalargleuileCysser 311
RESULT 32
US-10-174-588-348
; Sequence 348, Application US/10174588
; Publication No. US200300272661
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C28
; CURRENT APPLICATION NUMBER: US/10/174,588
; PRIOR FILING DATE: 2002-06-18
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-588-348
Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0
US-09-774-490-1 (1-2709) x US-10-174-588-348 (1-777)
QY 1040 AATAATGACAACTTCTCAAGCTCGTGTGATTGCTCA 1081
|||||
298 AenlystirpThrPheleuysalargleuileCysser 311
RESULT 33
US-10-175-739-348
; Sequence 348, Application US/10175739
; Publication No. US20030027267A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C46
; CURRENT APPLICATION NUMBER: US/10/175,739
; CURRENT FILING DATE: 2002-06-19

; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-739-348
Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0
US-09-774-490-1 (1-2709) x US-10-175-739-348 (1-777)
QY 1040 AATAATGACAACTTCTCAAGCTCGTGTGATTGCTCA 1081
|||||
298 AenlystirpThrPheleuysalargleuileCysser 311
RESULT 34
US-10-175-740-348
; Sequence 348, Application US/10175740
; Publication No. US20030027268A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C61
; CURRENT APPLICATION NUMBER: US/10/175,740
; PRIOR FILING DATE: 2002-06-18
; PRIOR APPLICATION removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-175-740-348
Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0
US-09-774-490-1 (1-2709) x US-10-175-740-348 (1-777)
QY 1040 AATAATGACAACTTCTCAAGCTCGTGTGATTGCTCA 1081
|||||
298 AenlystirpThrPheleuysalargleuileCysser 311
RESULT 35
US-10-175-743-348
; Sequence 348, Application US/10175743
; Publication No. US20030027269A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc

APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C52
CURRENT APPLICATION NUMBER: US/10/175, 743
CURRENT FILING DATE: 2002-06-16
PRIOR APPLICATION NUMBER: 10/052586
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059266
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063120
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063121
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063486
PRIOR FILING DATE: 1997-10-21
PRIOR APPLICATION NUMBER: 60/063540
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063541
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063544
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063870
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066120
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066466
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066772
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069335
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069425
PRIOR FILING DATE: 1997-12-12
PRIOR APPLICATION NUMBER: 60/069870
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/068017
PRIOR FILING DATE: 1997-12-18
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31

PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/08582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086023
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086486
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087098
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087208
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029

```
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088033
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088167
/ PRIOR FILING DATE: 1998-06-05
/ PRIOR APPLICATION NUMBER: 60/088202
/ PRIOR FILING DATE: 1998-06-05
/ PRIOR APPLICATION NUMBER: 60/088212
/ PRIOR FILING DATE: 1998-06-05
/ PRIOR APPLICATION NUMBER: 60/088217
/ PRIOR FILING DATE: 1998-06-05
/ PRIOR APPLICATION NUMBER: 60/088326
/ PRIOR FILING DATE: 1998-06-04
/ PRIOR APPLICATION NUMBER: 60/088655
/ PRIOR FILING DATE: 1998-06-09
/ PRIOR APPLICATION NUMBER: 60/088722
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088738
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088740
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088811
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088824
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088825
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088826
/ PRIOR FILING DATE: 1998-06-10
/ PRIOR APPLICATION NUMBER: 60/088861
/ PRIOR FILING DATE: 1998-06-11
/ PRIOR APPLICATION NUMBER: 60/088863
/ PRIOR FILING DATE: 1998-06-11
/ PRIOR APPLICATION NUMBER: 60/088876
/ PRIOR FILING DATE: 1998-06-11
/ PRIOR APPLICATION NUMBER: 60/089090
/ PRIOR FILING DATE: 1998-06-12
/ PRIOR APPLICATION NUMBER: 60/089105
/ PRIOR FILING DATE: 1998-06-12
/ PRIOR APPLICATION NUMBER: 60/089512
/ PRIOR FILING DATE: 1998-06-16
/ PRIOR APPLICATION NUMBER: 60/089514
/ PRIOR FILING DATE: 1998-06-16
/ PRIOR APPLICATION NUMBER: 60/089538
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089598
/ PRIOR FILING DATE: 1998-06-17
/ PRIOR APPLICATION NUMBER: 60/089653
```

```
Alignment Scores:
Pred. No.: 0.000106      length: 777
Score: 14.00             Matches: 14
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.56%              Indels: 0
DB: 15                       Gaps: 0
```

US-09-774-490-1 (1-2709) x US-10-176-743-348 (1-777)

```
OY 1040 AATAAATGACAACTTCCTCAAGCTGCTGATTGCTCA 1081
DB 238 AenlystipThrTheuleuylsalatrgleuileCySer 311
```

```
RESULT 36
US-10-176-488-348
/ Sequence 348, Application US/10176488
/ Publication No. US20030027271A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
```

```
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430R1C119
/ CURRENT APPLICATION NUMBER: US/10/176,488
/ PRIOR FILING DATE: 2002-06-21
/ PRIOR APPLICATION removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 348
/ LENGTH: 777
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-176-488-348
```

```
Alignment Scores:
Pred. No.: 0.000106      length: 777
Score: 14.00             Matches: 14
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.56%              Indels: 0
DB: 15                       Gaps: 0
```

US-09-774-490-1 (1-2709) x US-10-176-488-348 (1-777)

```
OY 1040 AATAAATGACAACTTCCTCAAGCTGCTGATTGCTCA 1081
DB 238 AenlystipThrTheuleuylsalatrgleuileCySer 311
```

```
RESULT 37
US-10-176-492-348
/ Sequence 348, Application US/10176492
/ Publication No. US20030027272A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430R1C107
/ CURRENT APPLICATION NUMBER: US/10/176,492
/ PRIOR FILING DATE: 2002-06-21
/ PRIOR APPLICATION removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 348
/ LENGTH: 777
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-176-492-348
```

```
Alignment Scores:
Pred. No.: 0.000106      length: 777
Score: 14.00             Matches: 14
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 1.56%              Indels: 0
DB: 15                       Gaps: 0
```

US-09-774-490-1 (1-2709) x US-10-176-492-348 (1-777)

```
OY 1040 AATAAATGACAACTTCCTCAAGCTGCTGATTGCTCA 1081
```

Db 298 AsnlystrpThrPheLeuYsAlaArgLeuIleCysSer 311

RESULT 38

US-10-176-747-348
; Sequence 348, Application US/10176747
; Publication No. US20030027278A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C92
; CURRENT APPLICATION NUMBER: US/10/176,747
; CURRENT FILING DATE: 2002-06-20
; Prior application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-747-348

Alignment Scores:

Pred. No.:	0.000106	Length:	777
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.56%	Indels:	0
DB:	15	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-176-747-348 (1-777)

QY 1040 AATAAATGACAACTTCCTCAAGCTCGTCGATTGCTCA 1081
Db 298 AsnlystrpThrPheLeuYsAlaArgLeuIleCysSer 311

RESULT 39

US-10-176-750-348
; Sequence 348, Application US/10176750
; Publication No. US20030027274A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C103
; CURRENT APPLICATION NUMBER: US/10/176,750
; CURRENT FILING DATE: 2002-06-21
; Prior application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien

US-10-176-750-348

Alignment Scores:

Pred. No.:	0.000106	Length:	777
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.56%	Indels:	0
DB:	15	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-176-750-348 (1-777)

QY 1040 AATAAATGACAACTTCCTCAAGCTCGTCGATTGCTCA 1081
Db 298 AsnlystrpThrPheLeuYsAlaArgLeuIleCysSer 311

RESULT 40

US-10-176-985-348
; Sequence 348, Application US/10176985
; Publication No. US20030027277A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C99
; CURRENT APPLICATION NUMBER: US/10/176,985
; CURRENT FILING DATE: 2002-06-21
; Prior application removed - See file wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-985-348

Alignment Scores:

Pred. No.:	0.000106	Length:	777
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.56%	Indels:	0
DB:	15	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-176-985-348 (1-777)

QY 1040 AATAAATGACAACTTCCTCAAGCTCGTCGATTGCTCA 1081
Db 298 AsnlystrpThrPheLeuYsAlaArgLeuIleCysSer 311

RESULT 41

US-10-176-987-348
; Sequence 348, Application US/10176987
; Publication No. US20030027278A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.

APPLICANT: Wood,William I.
APPLICANT: Zhang,Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C93
CURRENT FILING DATE: 2002-06-21
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-992-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-987-348 (1-777)

QY 1040 AATAATGACAACTTCTCAAGCTGCTGATTGCTCA 1081
|||||
Db 298 AenlystipthrThrPheleuylsAlaArgleuileCysSer 311

RESULT 42
US-10-176-992-348
Sequence 348, Application US/10176992
Publication No. US20030027279A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C100
CURRENT FILING DATE: 2002-06-21
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-992-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-992-348 (1-777)

QY 1040 AATAATGACAACTTCTCAAGCTGCTGATTGCTCA 1081
|||||
Db 298 AenlystipthrThrPheleuylsAlaArgleuileCysSer 311

RESULT 43

US-10-176-993-348
Sequence 348, Application US/10176993
Publication No. US20030027280A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C89
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-993-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-993-348 (1-777)

QY 1040 AATAATGACAACTTCTCAAGCTGCTGATTGCTCA 1081
|||||
Db 298 AenlystipthrThrPheleuylsAlaArgleuileCysSer 311

RESULT 44
US-10-184-658-348
Sequence 348, Application US/10184658
Publication No. US20030027281A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C228
CURRENT FILING DATE: 2002-06-28
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-184-658-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777

Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-184-658-348 (1-777)

QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
|||||
DB 298 AenlystiphrThrPheLeuYsAlaArgLeuileCySer 311

RESULT 45

US-10-176-991-348
; Sequence 348, Application US/10176991
; Publication No. US20030027324A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C122
; CURRENT APPLICATION NUMBER: US/10/176,991
; CURRENT FILING DATE: 2002-06-21
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-991-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-991-348 (1-777)

QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
|||||
DB 298 AenlystiphrThrPheLeuYsAlaArgLeuileCySer 311

RESULT 46

US-10-173-695-348
; Sequence 348, Application US/10173695
; Publication No. US20030032101A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C122
; CURRENT APPLICATION NUMBER: US/10/173,695
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-695-348

FILE REFERENCE: P3430R1C3
; CURRENT APPLICATION NUMBER: US/10/173,695
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-695-348

QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
|||||
DB 298 AenlystiphrThrPheLeuYsAlaArgLeuileCySer 311

RESULT 47

US-10-173-697-348
; Sequence 348, Application US/10173697
; Publication No. US20030032102A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C3
; CURRENT APPLICATION NUMBER: US/10/173,697
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-697-348

US-09-774-490-1 (1-2709) x US-10-173-695-348 (1-777)

QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
|||||
DB 298 AenlystiphrThrPheLeuYsAlaArgLeuileCySer 311

RESULT 48

US-10-173-705-348
; Sequence 348, Application US/10173705
; Publication No. US20030032103A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C3
; CURRENT APPLICATION NUMBER: US/10/173,705
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-705-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-173-697-348 (1-777)

QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081
|||||
DB 298 AenlystiphrThrPheLeuYsAlaArgLeuileCySer 311

RESULT 48

US-10-173-705-348
; Sequence 348, Application US/10173705
; Publication No. US20030032103A1
; GENERAL INFORMATION:

```

; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C18
; CURRENT APPLICATION NUMBER: US/10/173,705
; CURRENT FILING DATE: 2002-06-17
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-173-705-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-173-705-348 (1-777)
QY 1040 AATAAATGACACATTCTCAAGCTGCTGATTGCTCA 1081
DB 298 AenlystptntrthPheleuylsAlargleuileCysser 311

RESULT 49
US-10-174-576-348
; Sequence 348, Application US/10174576
; Publication No. US20030032104A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C23
; CURRENT APPLICATION NUMBER: US/10/174,576
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-576-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0
```

```

DB: 15 Gaps: 0
US-09-774-490-1 (1-2709) x US-10-174-576-348 (1-777)
QY 1040 AATAAATGACACATTCTCAAGCTGCTGATTGCTCA 1081
DB 298 AenlystptntrthPheleuylsAlargleuileCysser 311

RESULT 50
US-10-174-585-348
; Sequence 348, Application US/10174585
; Publication No. US20030032105A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C17
; CURRENT APPLICATION NUMBER: US/10/174,585
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-174-585-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-174-585-348 (1-777)
QY 1040 AATAAATGACACATTCTCAAGCTGCTGATTGCTCA 1081
DB 298 AenlystptntrthPheleuylsAlargleuileCysser 311

RESULT 51
US-10-174-586-348
; Sequence 348, Application US/10174586
; Publication No. US20030032106A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3430R1C24
; CURRENT APPLICATION NUMBER: US/10/174,586
; CURRENT FILING DATE: 2002-06-18
; Prior Application removed - See File Wrapper or Palm
```

```
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 348
/ LENGTH: 777
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-174-586-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-174-586-348 (1-777)

QY 1040 AATAAATGACAACTTCTCAAGCTGCTGATTTGCTCA 1081
    |||||
DB 298 Asnlystrphtthrphleuylsalargleuilecysser 311

RESULT 52
US-10-175-747-348
/ Sequence 348, Application US/10175747
/ Publication No. US20030032107A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Matanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430R1C44
/ CURRENT APPLICATION NUMBER: US/10/175,747
/ CURRENT FILING DATE: 2002-06-19
/ PRIOR APPLICATION NUMBER: 10/052586
/ PRIOR FILING DATE: 2002-01-15
/ PRIOR APPLICATION NUMBER: 60/059263
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/059266
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/062250
/ PRIOR FILING DATE: 1997-10-17
/ PRIOR APPLICATION NUMBER: 60/063120
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/063121
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/063486
/ PRIOR FILING DATE: 1997-10-21
/ PRIOR APPLICATION NUMBER: 60/063540
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063541
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063544
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063564
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063734
/ PRIOR FILING DATE: 1997-10-29
/ PRIOR APPLICATION NUMBER: 60/063870
/ PRIOR FILING DATE: 1997-10-31
/ PRIOR APPLICATION NUMBER: 60/064103
/ PRIOR FILING DATE: 1997-10-31
/ PRIOR APPLICATION NUMBER: 60/065311
/ PRIOR FILING DATE: 1997-11-13
/ PRIOR APPLICATION NUMBER: 60/066120
```

```
/ PRIOR FILING DATE: 1997-11-21
/ PRIOR APPLICATION NUMBER: 60/066466
/ PRIOR FILING DATE: 1997-11-24
/ PRIOR APPLICATION NUMBER: 60/066772
/ PRIOR FILING DATE: 1997-11-24
/ PRIOR APPLICATION NUMBER: 60/069335
/ PRIOR FILING DATE: 1997-12-11
/ PRIOR APPLICATION NUMBER: 60/069425
/ PRIOR FILING DATE: 1997-12-12
/ PRIOR APPLICATION NUMBER: 60/069870
/ PRIOR FILING DATE: 1997-12-17
/ PRIOR APPLICATION NUMBER: 60/068017
/ PRIOR FILING DATE: 1997-12-18
/ PRIOR APPLICATION NUMBER: 60/077450
/ PRIOR FILING DATE: 1998-03-10
/ PRIOR APPLICATION NUMBER: 60/077632
/ PRIOR FILING DATE: 1998-03-11
/ PRIOR APPLICATION NUMBER: 60/077649
/ PRIOR FILING DATE: 1998-03-11
/ PRIOR APPLICATION NUMBER: 60/078886
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/078939
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/079664
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: 60/079786
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: 60/080107
/ PRIOR FILING DATE: 1998-03-31
/ PRIOR APPLICATION NUMBER: 60/080194
/ PRIOR FILING DATE: 1998-03-31
/ PRIOR APPLICATION NUMBER: 60/080327
/ PRIOR FILING DATE: 1998-04-01
/ PRIOR APPLICATION NUMBER: 60/080333
/ PRIOR FILING DATE: 1998-04-01
/ PRIOR APPLICATION NUMBER: 60/081049
/ PRIOR FILING DATE: 1998-04-08
/ PRIOR APPLICATION NUMBER: 60/081070
/ PRIOR FILING DATE: 1998-04-08
/ PRIOR APPLICATION NUMBER: 60/081195
/ PRIOR FILING DATE: 1998-04-09
/ PRIOR APPLICATION NUMBER: 60/081838
/ PRIOR FILING DATE: 1998-04-15
/ PRIOR APPLICATION NUMBER: 60/082568
/ PRIOR FILING DATE: 1998-04-21
/ PRIOR APPLICATION NUMBER: 60/082569
/ PRIOR FILING DATE: 1998-04-21
/ PRIOR APPLICATION NUMBER: 60/082704
/ PRIOR FILING DATE: 1998-04-22
/ PRIOR APPLICATION NUMBER: 60/082797
/ PRIOR FILING DATE: 1998-04-22
/ PRIOR APPLICATION NUMBER: 60/083322
/ PRIOR FILING DATE: 1998-04-28
/ PRIOR APPLICATION NUMBER: 60/083495
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083496
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083499
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083559
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/084366
/ PRIOR FILING DATE: 1998-05-05
/ PRIOR APPLICATION NUMBER: 60/084414
/ PRIOR FILING DATE: 1998-05-06
/ PRIOR APPLICATION NUMBER: 60/084639
/ PRIOR FILING DATE: 1998-05-07
/ PRIOR APPLICATION NUMBER: 60/084640
/ PRIOR FILING DATE: 1998-05-07
/ PRIOR APPLICATION NUMBER: 60/084643
/ PRIOR FILING DATE: 1998-05-07
/ PRIOR APPLICATION NUMBER: 60/085573
/ PRIOR FILING DATE: 1998-05-15
```


PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086023
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086486
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087098
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087208
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088722
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088740
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088811
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088825
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088863
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089090
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089538

PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653

Alignment Scores:

Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-747-348 (1-777)

QY 1040 AATAAATGACAACTTCCTCAAGCTGCTGATTGCTCA 1081
|||||
Db 298 AenlySTpThrThrPheleuylsAlaArgleuileCysSer 311

RESULT 53

US-10-176-481-348
Sequence 348, Application US/10176481
Publication No. US20030032108A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.

APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3430R1C98
CURRENT APPLICATION NUMBER: US/10/176,481
CURRENT FILING DATE: 2002-06-21
Prior application removed - See file Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348

LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-481-348

Alignment Scores:

Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-481-348 (1-777)

QY 1040 AATAAATGACAACTTCCTCAAGCTGCTGATTGCTCA 1081
|||||
Db 298 AenlySTpThrThrPheleuylsAlaArgleuileCysSer 311

RESULT 54

US-10-176-485-348
Sequence 348, Application US/10176485
Publication No. US20030032109A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.

APPLICANT: Pan,James
APPLICANT: Smith,Victoria
APPLICANT: Watanabe,Colin K.
APPLICANT: Wood,William I.
APPLICANT: Zhang,Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C78
CURRENT APPLICATION NUMBER: US/10/176,485
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-485-348

Alignment Scores:

Pred. No.:	0.000106	Length:	777
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.56%	Indels:	0
DB:	15	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-176-485-348 (1-777)

QY 1040 AATAATGACAACTTCTCAAGCTGCTGATTGCTCA 1081
|||||
Db 298 AenlystrpThrThrpheulysAlaArgleuileCySer 311
|||||

RESULT 55

US-10-176-487-348
Sequence 348, Application US/10176487
Publication No. US20030032110A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C74
CURRENT APPLICATION NUMBER: US/10/176,487
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-487-348

Alignment Scores:

Pred. No.:	0.000106	Length:	777
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.56%	Indels:	0
DB:	15	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-176-487-348 (1-777)

QY 1040 AATAATGACAACTTCTCAAGCTGCTGATTGCTCA 1081
|||||
|||||

Db 298 AenlystrpThrThrpheulysAlaArgleuileCySer 311
|||||

RESULT 56

US-10-176-493-348
Sequence 348, Application US/10176493
Publication No. US20030032111A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C72
CURRENT APPLICATION NUMBER: US/10/176,493
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-493-348

Alignment Scores:

Pred. No.:	0.000106	Length:	777
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.56%	Indels:	0
DB:	15	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-176-493-348 (1-777)

QY 1040 AATAATGACAACTTCTCAAGCTGCTGATTGCTCA 1081
|||||
Db 298 AenlystrpThrThrpheulysAlaArgleuileCySer 311
|||||

RESULT 57

US-10-176-756-348
Sequence 348, Application US/10176756
Publication No. US20030032112A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C109
CURRENT APPLICATION NUMBER: US/10/176,756
CURRENT FILING DATE: 2002-06-21
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-756-348

QY 1040 AATAATGACAACTTCTCAAGCTGCTGATTGCTCA 1081
|||||
|||||


```
/ Sequence 348, Application US/10176378
/ Publication No. US20030032116A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430R1C116
/ CURRENT FILING DATE: 2002-06-21
/ Prior application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 348
/ LENGTH: 777
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-176-978-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-176-978-348 (1-777)
QY 1040 AATAATGACAACTTCCTCAAGCTGCTGATTGCTCA 1081
Db 298 AenlystrpThrPheLeuYsaIaArgLeuIleCySer 311

RESULT 62
US-10-179-510-348
/ Sequence 348, Application US/10179510
/ Publication No. US20030032117A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430R1C138
/ CURRENT FILING DATE: 2002-06-24
/ Prior application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 348
/ LENGTH: 777
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-179-510-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0
```

```
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-179-510-348 (1-777)
QY 1040 AATAATGACAACTTCCTCAAGCTGCTGATTGCTCA 1081
Db 298 AenlystrpThrPheLeuYsaIaArgLeuIleCySer 311

RESULT 63
US-10-180-543-348
/ Sequence 348, Application US/10180543
/ Publication No. US20030032118A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430R1C161
/ CURRENT FILING DATE: 2002-06-25
/ Prior application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 348
/ LENGTH: 777
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-180-543-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-180-543-348 (1-777)
QY 1040 AATAATGACAACTTCCTCAAGCTGCTGATTGCTCA 1081
Db 298 AenlystrpThrPheLeuYsaIaArgLeuIleCySer 311

RESULT 64
US-10-180-544-348
/ Sequence 348, Application US/10180544
/ Publication No. US20030032119A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430R1C150
```

```

; CURRENT APPLICATION NUMBER: US/10/180,544
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-544-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-180-544-348 (1-777)
QY 1040 AATAATGACACACTTCTCAAGCTGCTGATTGCTCA 1081
DB 298 AsnysrtprrThrPheleuysAlaArgleuileCysSer 311

RESULT 65
US-10-180-546-348
; Sequence 348, Application US/10180546
; Publication No. US20030032120A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C156
; CURRENT APPLICATION NUMBER: US/10/180,546
; CURRENT FILING DATE: 2002-06-25
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-546-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-180-546-348 (1-777)
QY 1040 AATAATGACACACTTCTCAAGCTGCTGATTGCTCA 1081
DB 298 AsnysrtprrThrPheleuysAlaArgleuileCysSer 311

RESULT 66
US-10-180-547-348
; Sequence 348, Application US/10180547
; Publication No. US20030032121A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

```

```

; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C157
; CURRENT APPLICATION NUMBER: US/10/180,547
; CURRENT FILING DATE: 2002-06-25
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-547-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-180-547-348 (1-777)
QY 1040 AATAATGACACACTTCTCAAGCTGCTGATTGCTCA 1081
DB 298 AsnysrtprrThrPheleuysAlaArgleuileCysSer 311

RESULT 67
US-10-180-549-348
; Sequence 348, Application US/10180549
; Publication No. US20030032122A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Chen, Jian
; APPLICANT: Desnoyers, Luc
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Smith, Victoria
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3430R1C151
; CURRENT APPLICATION NUMBER: US/10/180,549
; CURRENT FILING DATE: 2002-06-25
; Prior application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 612
; SEQ ID NO 348
; LENGTH: 777
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-180-549-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

```

US-09-774-490-1 (1-2709) x US-10-180-549-348 (1-777)

QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081

DB 298 AenlystipThrThrPheLeuLysAlaArgLeuIleCysSer 311

RESULT 68

US-10-180-555-348

/ Sequence 348, Application US/10180555

/ Publication No. US20030032123A1

/ GENERAL INFORMATION:

/ APPLICANT: Baker, Kevin P.

/ APPLICANT: Chen, Jian

/ APPLICANT: Desnoyers, Luc

/ APPLICANT: Goddard, Audrey

/ APPLICANT: Godowski, Paul J.

/ APPLICANT: Gurney, Austin L.

/ APPLICANT: Pan, James

/ APPLICANT: Smith, Victoria

/ APPLICANT: Watanabe, Colin K.

/ APPLICANT: Wood, William I.

/ APPLICANT: Zhang, Zemin

/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

/ FILE REFERENCE: P3430R1C163

/ CURRENT APPLICATION NUMBER: US/10/180,555

/ CURRENT FILING DATE: 2002-06-25

/ Prior Application removed - See File Wrapper or Palm

/ NUMBER OF SEQ ID NOS: 612

/ SEQ ID NO 348

/ LENGTH: 777

/ TYPE: PRT

/ ORGANISM: Homo Sapien

US-10-180-555-348

Alignment Scores:

Pred. No.: 0.000106 Length: 777

Score: 14.00 Matches: 14

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.56% Indels: 0

DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-180-555-348 (1-777)

QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081

DB 298 AenlystipThrThrPheLeuLysAlaArgLeuIleCysSer 311

RESULT 69

US-10-180-559-348

/ Sequence 348, Application US/10180559

/ Publication No. US20030032124A1

/ GENERAL INFORMATION:

/ APPLICANT: Baker, Kevin P.

/ APPLICANT: Chen, Jian

/ APPLICANT: Desnoyers, Luc

/ APPLICANT: Goddard, Audrey

/ APPLICANT: Godowski, Paul J.

/ APPLICANT: Gurney, Austin L.

/ APPLICANT: Pan, James

/ APPLICANT: Smith, Victoria

/ APPLICANT: Watanabe, Colin K.

/ APPLICANT: Wood, William I.

/ APPLICANT: Zhang, Zemin

/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

/ FILE REFERENCE: P3430R1C159

/ CURRENT APPLICATION NUMBER: US/10/180,559

/ CURRENT FILING DATE: 2002-06-25

/ Prior Application removed - See File Wrapper or Palm

/ NUMBER OF SEQ ID NOS: 612

/ SEQ ID NO 348

/ LENGTH: 777

/ TYPE: PRT

/ ORGANISM: Homo Sapien

US-10-180-559-348

Alignment Scores:

Pred. No.: 0.000106 Length: 777

Score: 14.00 Matches: 14

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.56% Indels: 0

DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-180-559-348 (1-777)

QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081

DB 298 AenlystipThrThrPheLeuLysAlaArgLeuIleCysSer 311

RESULT 70

US-10-181-000-348

/ Sequence 348, Application US/10181000

/ Publication No. US20030032125A1

/ GENERAL INFORMATION:

/ APPLICANT: Baker, Kevin P.

/ APPLICANT: Chen, Jian

/ APPLICANT: Desnoyers, Luc

/ APPLICANT: Goddard, Audrey

/ APPLICANT: Godowski, Paul J.

/ APPLICANT: Gurney, Austin L.

/ APPLICANT: Pan, James

/ APPLICANT: Smith, Victoria

/ APPLICANT: Watanabe, Colin K.

/ APPLICANT: Wood, William I.

/ APPLICANT: Zhang, Zemin

/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

/ FILE REFERENCE: P3430R1C177

/ CURRENT APPLICATION NUMBER: US/10/181,000

/ CURRENT FILING DATE: 2002-06-26

/ Prior Application removed - See File Wrapper or Palm

/ NUMBER OF SEQ ID NOS: 612

/ SEQ ID NO 348

/ LENGTH: 777

/ TYPE: PRT

/ ORGANISM: Homo Sapien

US-10-181-000-348

Alignment Scores:

Pred. No.: 0.000106 Length: 777

Score: 14.00 Matches: 14

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.56% Indels: 0

DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-181-000-348 (1-777)

QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081

DB 298 AenlystipThrThrPheLeuLysAlaArgLeuIleCysSer 311

RESULT 71

US-10-183-010-348

/ Sequence 348, Application US/10183010

/ Publication No. US20030032126A1

/ GENERAL INFORMATION:

/ APPLICANT: Baker, Kevin P.

/ APPLICANT: Chen, Jian

/ APPLICANT: Desnoyers, Luc

/ APPLICANT: Goddard, Audrey

/ APPLICANT: Godowski, Paul J.

```

/ APPLICANT: Gurney,Austin L.
/ APPLICANT: Pan,James
/ APPLICANT: Smith,Victoria
/ APPLICANT: Watanabe,Colin K.
/ APPLICANT: Wood,William I.
/ APPLICANT: Zhang,Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430R1C164
/ CURRENT APPLICATION NUMBER: US/10/183,010
/ CURRENT FILING DATE: 2002-06-26
/ Prior Application removed - See File Wrapper or Palm
/ NUMBER OF SEQ ID NOS: 612
/ SEQ ID NO 348
/ LENGTH: 777
/ TYPE: PRT
/ ORGANISM: Homo Sapien
US-10-183-010-348

Alignment Scores:
Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-183-010-348 (1-777)
QY 1040 AATTAATGACAACTTCTCAAGCTGCTGATTTGCTCA 1081
DB 298 AasnlystipThrPhreheutysalaxgleuilecyser 311

RESULT 72
US-10-183-012-348
/ Sequence 348, Application US/10183012
/ Publication No. US20030032127A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Chen, Jian
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Pan, James
/ APPLICANT: Smith, Victoria
/ APPLICANT: Watanabe, Colin K.
/ APPLICANT: Wood, William I.
/ APPLICANT: Zhang, Zemin
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ FILE REFERENCE: P3430R1C171
/ CURRENT APPLICATION NUMBER: US/10/183,012
/ CURRENT FILING DATE: 2002-06-26
/ PRIOR APPLICATION NUMBER: 10/052586
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/059263
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/059266
/ PRIOR FILING DATE: 1997-09-18
/ PRIOR APPLICATION NUMBER: 60/062250
/ PRIOR FILING DATE: 1997-10-17
/ PRIOR APPLICATION NUMBER: 60/063120
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/063121
/ PRIOR FILING DATE: 1997-10-24
/ PRIOR APPLICATION NUMBER: 60/063486
/ PRIOR FILING DATE: 1997-10-21
/ PRIOR APPLICATION NUMBER: 60/063540
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063541
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063544
```

```

/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063564
/ PRIOR FILING DATE: 1997-10-28
/ PRIOR APPLICATION NUMBER: 60/063734
/ PRIOR FILING DATE: 1997-10-29
/ PRIOR APPLICATION NUMBER: 60/063870
/ PRIOR FILING DATE: 1997-10-31
/ PRIOR APPLICATION NUMBER: 60/064103
/ PRIOR FILING DATE: 1997-10-31
/ PRIOR APPLICATION NUMBER: 60/065311
/ PRIOR FILING DATE: 1997-11-13
/ PRIOR APPLICATION NUMBER: 60/066120
/ PRIOR FILING DATE: 1997-11-21
/ PRIOR APPLICATION NUMBER: 60/066466
/ PRIOR FILING DATE: 1997-11-24
/ PRIOR APPLICATION NUMBER: 60/066772
/ PRIOR FILING DATE: 1997-11-24
/ PRIOR APPLICATION NUMBER: 60/069335
/ PRIOR FILING DATE: 1997-12-11
/ PRIOR APPLICATION NUMBER: 60/069425
/ PRIOR FILING DATE: 1997-12-12
/ PRIOR APPLICATION NUMBER: 60/069870
/ PRIOR FILING DATE: 1997-12-17
/ PRIOR APPLICATION NUMBER: 60/068017
/ PRIOR FILING DATE: 1997-12-18
/ PRIOR APPLICATION NUMBER: 60/07450
/ PRIOR FILING DATE: 1998-03-10
/ PRIOR APPLICATION NUMBER: 60/077632
/ PRIOR FILING DATE: 1998-03-11
/ PRIOR APPLICATION NUMBER: 60/077649
/ PRIOR FILING DATE: 1998-03-11
/ PRIOR APPLICATION NUMBER: 60/078886
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/078939
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/079664
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: 60/079786
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: 60/080107
/ PRIOR FILING DATE: 1998-03-31
/ PRIOR APPLICATION NUMBER: 60/080194
/ PRIOR FILING DATE: 1998-03-31
/ PRIOR APPLICATION NUMBER: 60/080327
/ PRIOR FILING DATE: 1998-04-01
/ PRIOR APPLICATION NUMBER: 60/080333
/ PRIOR FILING DATE: 1998-04-01
/ PRIOR APPLICATION NUMBER: 60/081049
/ PRIOR FILING DATE: 1998-04-08
/ PRIOR APPLICATION NUMBER: 60/081070
/ PRIOR FILING DATE: 1998-04-08
/ PRIOR APPLICATION NUMBER: 60/081195
/ PRIOR FILING DATE: 1998-04-09
/ PRIOR APPLICATION NUMBER: 60/081838
/ PRIOR FILING DATE: 1998-04-15
/ PRIOR APPLICATION NUMBER: 60/082568
/ PRIOR FILING DATE: 1998-04-21
/ PRIOR APPLICATION NUMBER: 60/082569
/ PRIOR FILING DATE: 1998-04-21
/ PRIOR APPLICATION NUMBER: 60/082704
/ PRIOR FILING DATE: 1998-04-22
/ PRIOR APPLICATION NUMBER: 60/082797
/ PRIOR FILING DATE: 1998-04-22
/ PRIOR APPLICATION NUMBER: 60/083322
/ PRIOR FILING DATE: 1998-04-28
/ PRIOR APPLICATION NUMBER: 60/083495
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083496
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083499
/ PRIOR FILING DATE: 1998-04-29
/ PRIOR APPLICATION NUMBER: 60/083559
/ PRIOR FILING DATE: 1998-04-29
```

```

PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086023
PRIOR FILING DATE: 1998-05-18
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086486
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087098
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087208
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/087609
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087759
PRIOR FILING DATE: 1998-06-02
PRIOR APPLICATION NUMBER: 60/087827
PRIOR FILING DATE: 1998-06-03
PRIOR APPLICATION NUMBER: 60/088025
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088655
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088722
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088728
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088740
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088811
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088825
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088863

```

```

PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089090
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653

```

Alignment Scores:

```

Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

```

US-09-774-490-1 (1-2709) x US-10-183-012-348 (1-777)

QY 1040 AATTAATGACAACTTCTCAAGCTGCTGATTGCTCA 1081

DB 298 AaLlystIpThrThPheLyuLsAlaArgLeuIleCySer 311

RESULT 73

```

US-10-184-614-348
Sequence 348, Application US/10184614
Publication No. US20030032128A1
GENERAL INFORMATION:

```

```

APPLICANT: Baker, Kevin P.
APPLICANT: Chen, Jian
APPLICANT: Desnoyers, Luc
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Pan, James
APPLICANT: Smith, Victoria
APPLICANT: Watanabe, Colin K.
APPLICANT: Wood, William I.
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3430R1C184
CURRENT APPLICATION NUMBER: US/10/184,614
CURRENT FILING DATE: 2225-06-27
Prior Application removed - See File Wrapper or Palm
Number of SEQ ID NOS: 612
SEQ ID NO 348
LENGTH: 777
TYPE: PRT
ORGANISM: Homo Sapien
US-10-184-614-348

```

Alignment Scores:

```

Pred. No.: 0.000106 Length: 777
Score: 14.00 Matches: 14
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.56% Indels: 0
DB: 15 Gaps: 0

```

US-09-774-490-1 (1-2709) x US-10-184-614-348 (1-777)

QY 1040 AATTAATGACAACTTCTCAAGCTGCTGATTGCTCA 1081

DB 298 AaLlystIpThrThPheLyuLsAlaArgLeuIleCySer 311

RESULT 74 f

US-10-184-623-348

; Sequence 348, Application US/10184623

; Publication No. US20030032129A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C210

; CURRENT APPLICATION NUMBER: US/10/184,623

; CURRENT FILING DATE: 2002-06-28

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 348

; LENGTH: 777

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-184-623-348

Alignment Scores:

Pred. No.:	0.000106	Length:	777
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.56%	Indels:	0
DB:	15	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-184-623-348 (1-777)

QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081

DB 298 AenlystiprthrPheleuylsAlaArgLeuileCysSer 311

RESULT 75

US-10-184-635-348

; Sequence 348, Application US/10184635

; Publication No. US20030032130A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Chen, Jian

; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Smith, Victoria

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

; APPLICANT: Zhang, Zemin

; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

; FILE REFERENCE: P3430R1C215

; CURRENT APPLICATION NUMBER: US/10/184,635

; CURRENT FILING DATE: 2002-06-28

; Prior Application removed - See File Wrapper or Palm

; NUMBER OF SEQ ID NOS: 612

; SEQ ID NO 348

; LENGTH: 777

; TYPE: PRT

; ORGANISM: Homo Sapien

US-10-184-635-348

Alignment Scores:

Pred. No.:	0.000106	Length:	777
Score:	14.00	Matches:	14
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.56%	Indels:	0
DB:	15	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-184-635-348 (1-777)

QY 1040 AATAATGACACATTCCTCAAGCTGCTGATTGCTCA 1081

DB 298 AenlystiprthrPheleuylsAlaArgLeuileCysSer 311

Search completed: August 3, 2003, 10:58:20
Job time : 124 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_r2p model

Run on: August 3, 2003, 10:29:36 ; Search time 401 Seconds

(without alignments)
11759.598 Million cell updates/sec

Title: US-09-774-490-1
Perfect score: 895
Sequence: 1 aactcttatttcacgacg.....agccttttcccaataacc 2709

Scoring table: OLIGO
Xgapop 60.0, Xgapext 60.0
Ygapop 60.0, Ygapext 60.0
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 5580241 seqs, 870357830 residues

Word size: 1

Total number of hits satisfying chosen parameters: 10777494

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 100 summaries

Command line parameters:

-MODEL=frame+r2p_model -DEV=xlh
-O=/cgn2_1/USFPO/seqs/US09774490/runat_03082003.102928.4506/app_query.fasta.1.2887
-DB=Pending_Patents_AA_Main -QMT=fastan -SUFFIX=oligo.rapm -MINMATCH=0.1
-LOOPL=0 -LOOPEXT=0 -UNITS=DITS -START=1 -END=1 -MATRIX=oligo
-ALIGN=human40.cdi -LIST=100 -DOCALIGN=200 -THR SCORE=quality -THR MIN=1
-MAXLEN=200000000 -USER=US09774490 @CGN 1.1 628 @runat_03082003.102928.4506
-NCPU=6 -ICPU=3 -NO_MMAP -LANG=ENGLISH -NEG_SCORES=50 -WAIT -DSPBLOCK=100
-LONGLOG -DBTIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Pending Patents_AA_Main:*

1: /cgn2_6/ptodata/2/paa/PCTUS_COMB.pep:*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pep:*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pep:*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pep:*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pep:*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pep:*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pep:*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pep:*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pep:*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pep:*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pep:*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pep:*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pep:*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pep:*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pep:*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pep:*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pep:*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pep:*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pep:*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pep:*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pep:*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pep:*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pep:*
24: /cgn2_6/ptodata/2/paa/US099A_COMB.pep:*
25: /cgn2_6/ptodata/2/paa/US099B_COMB.pep:*
26: /cgn2_6/ptodata/2/paa/US100_COMB.pep:*
27: /cgn2_6/ptodata/2/paa/US101_COMB.pep:*
28: /cgn2_6/ptodata/2/paa/US102_COMB.pep:*

29: /cgn2_6/ptodata/2/paa/US103_COMB.pep:*
30: /cgn2_6/ptodata/2/paa/US104_COMB.pep:*
31: /cgn2_6/ptodata/2/paa/US60_COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	* Query Match	Length	DB ID	Description
1	771	86.1	771	1 PCT-US00-41943-24	Sequence 24, Appl
2	771	86.1	771	1 PCT-US02-07926-284	Sequence 24, Appl
3	771	86.1	771	8 US-08-488-935-2	Sequence 2, Appl
4	771	86.1	771	8 US-08-488-935-2	Sequence 2, Appl
5	771	86.1	771	20 US-09-694-085-24	Sequence 24, Appl
6	771	86.1	771	22 US-09-791-537-31106	Sequence 31106, A
7	771	86.1	771	26 US-10-067-632-54	Sequence 54, Appl
8	771	86.1	771	26 US-10-097-340-284	Sequence 284, Appl
9	771	86.1	771	28 US-10-247-671-164	Sequence 164, Appl
10	771	86.1	771	28 US-10-262-538-10	Sequence 10, Appl
11	771	86.1	771	31 US-09-791-537-32993	Sequence 164, Appl
12	771	86.1	771	31 US-09-791-537-32993	Sequence 32993, A
13	771	86.1	771	21 US-09-724-676-66689	Sequence 66689, A
14	771	86.1	771	21 US-09-724-676-66689	Sequence 66689, A
15	771	86.1	771	23 US-09-864-761-47112	Sequence 47112, A
16	771	86.1	771	28 US-10-203-135-32555	Sequence 32555, A
17	771	86.1	771	28 US-10-203-136-33448	Sequence 33448, A
18	771	86.1	771	22 US-09-791-537-101922	Sequence 101922, A
19	771	86.1	771	22 US-09-791-537-101922	Sequence 62757, A
20	771	86.1	771	16 US-09-791-537-62757	Sequence 32, Appl
21	771	86.1	771	22 US-09-791-537-62757	Sequence 128280, A
22	771	86.1	771	22 US-09-791-537-62757	Sequence 3560, Ap
23	771	86.1	771	31 US-09-791-537-62757	Sequence 4370, Ap
24	771	86.1	771	31 US-09-791-537-62757	Sequence 111919, A
25	771	86.1	771	31 US-09-791-537-62757	Sequence 111919, A
26	771	86.1	771	31 US-09-791-537-62757	Sequence 30748, A
27	771	86.1	771	26 US-10-029-186-30748	Sequence 33541, A
28	771	86.1	771	1 PCT-US01-00663-33648	Sequence 33648, A
29	771	86.1	771	57 US-09-864-761-44552	Sequence 44552, A
30	771	86.1	771	27 US-10-182-993-32519	Sequence 32519, A
31	771	86.1	771	27 US-10-182-993-32519	Sequence 33556, A
32	771	86.1	771	28 US-10-203-135-32559	Sequence 32559, A
33	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
34	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
35	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
36	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
37	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
38	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
39	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
40	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
41	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
42	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
43	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
44	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
45	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
46	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
47	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
48	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
49	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
50	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
51	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
52	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
53	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
54	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
55	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
56	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
57	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
58	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
59	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A
60	771	86.1	771	28 US-10-203-135-32559	Sequence 33559, A

```

61 17 1.9 17 1 PCT-US00-41943-16 Sequence 16, Appl
62 17 1.9 17 1 PCT-US00-41943-43 Sequence 43, Appl
63 17 1.9 17 20 US-09-694-085-16 Sequence 16, Appl
64 17 1.9 17 20 US-09-694-085-43 Sequence 43, Appl
65 17 1.9 73 31 US-60-177-571-3765 Sequence 3765, Ap
66 17 1.9 73 31 US-60-177-571-3105 Sequence 3105, Ap
67 17 1.9 79 31 US-60-177-571-3766 Sequence 3766, Ap
68 17 1.9 779 18 PCT-US00-33116-2 Sequence 2, Appl
69 17 1.9 779 21 US-09-455-560-2 Sequence 2, Appl
70 17 1.9 779 21 US-09-731-179-2 Sequence 2, Appl
71 17 1.9 782 19 PCT-US00-04340-18 Sequence 18, Appl
72 17 1.9 782 23 US-09-507-209-18 Sequence 18, Appl
73 17 1.9 782 23 US-09-813-290-4 Sequence 4, Appl
74 17 1.9 785 19 PCT-US00-04340-92 Sequence 92, Appl
75 17 1.9 785 19 US-09-507-209-92 Sequence 92, Appl
76 17 1.9 875 23 US-09-813-290-2 Sequence 2, Appl
77 17 1.8 17 1 PCT-US00-41943-2 Sequence 2, Appl
78 17 1.8 17 20 US-09-694-085-2 Sequence 2, Appl
79 17 1.8 34 31 US-60-177-571-3504 Sequence 3504, Ap
80 17 1.8 764 22 US-09-791-537-85681 Sequence 85681, A
81 17 1.7 39 31 US-60-170-373-3437 Sequence 3437, Ap
82 17 1.7 89 31 US-60-196-718-4235 Sequence 4235, Ap
83 17 1.7 110 31 US-60-177-571-4486 Sequence 4486, Ap
84 17 1.7 478 1 PCT-US02-40059-15 Sequence 15, Appl
85 17 1.7 539 22 US-09-760-479-723 Sequence 723, App
86 17 1.7 539 28 US-10-206-008-723 Sequence 723, App
87 17 1.7 539 28 US-10-211-364-1199 Sequence 1199, Ap
88 17 1.7 540 22 US-09-760-479-483 Sequence 483, App
89 17 1.7 540 28 US-10-206-008-483 Sequence 483, App
90 17 1.7 636 31 US-60-389-987-2003 Sequence 2003, Ap
91 17 1.7 636 31 US-60-412-418-2003 Sequence 2003, Ap
92 17 1.7 658 31 US-60-242-679-840 Sequence 840, App
93 17 1.7 691 16 US-09-262-167-33 Sequence 33, Appl
94 17 1.7 701 16 US-09-262-167-29 Sequence 29, Appl
95 17 1.7 749 1 PCT-US00-41943-25 Sequence 25, Appl
96 17 1.7 749 20 US-09-694-085-25 Sequence 25, Appl
97 17 1.7 749 28 US-10-262-538-12 Sequence 12, Appl
98 17 1.7 749 28 US-10-285-3518-1 Sequence 1, Appl
99 17 1.7 749 31 US-60-335-783-1 Sequence 4, Appl
100 15 1.7 751 1 PCT-US00-10046-4 Sequence 4, Appl

```

ALIGNMENTS

```

RESULT 1
PCT-US00-41943-24
; Sequence 24, Application PC/TUS0041943
; GENERAL INFORMATION:
; APPLICANT: The General Hospital Corporation
; TITLE OF INVENTION: NOVEL SEMAPHORIN DOMAINS
; FILE REFERENCE: 00786-441M01
; CURRENT APPLICATION NUMBER: PCT/US00/41943
; CURRENT FILING DATE: 2000-11-07
; PRIOR APPLICATION NUMBER: US 09/694, 085
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/164, 056
; PRIOR FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US00-41943-24

```

Alignment Scores:

```

Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 1 Gaps: 0

```

```

US-09-774-490-1 (1-2709) x PCT-US00-41943-24 (1-771)
QY 200 ATGGGCTGGTTAACTAGATTTGCTGCTTTCTGGGAGTATTACTTACAGCAAGACA 259
Db 1 MetGlyTyrLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrIlaArgAla 20
QY 260 AACCTGACAAATGGGAGAGACAATGTCGCAAGCTGTAATTTCTTACAAAGAAATGTTG 319
Db 21 AenTyrGlnAsnGlyValAsnAsnValProAlaGlyLeuValSerTyrIleGlyMetLeu 40
QY 320 GAATCAACAATGATGATCATCTTCAATGCTTGGCCAAAGCTCCAGTTATCATACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuIlaAsnSerSerTyrIleThrPhe 60
QY 380 CTTTGGAGAGAGAACGAGTGGCTGTATGTTGGAGCAAGAGATCACAATTTTCATTC 439
Db 61 LeuLeuAspGlyLysArgSerArgLeuTyrValGlyAlaValAspHisIlePheSerPhe 80
QY 440 GACCTGTTAATATATCAAGGATTTTCAAAAGATTGTCGCGCAGTATCTTACACCAAGA 499
Db 81 AspLeuValAsnIleLeuAspPheGlnIleValIleProValSerTyrThrArgArg 100
QY 500 GATGAATGCAAGTGGCGCTGGAAAGACATCTGAAAGAATGTGTTATTCATCAAGTA 559
Db 101 AspGlyLysTyrTrpAlaGlyLysAspIleLeuLysGlyLysAlaAsnPheIleVal 120
QY 560 CTTAAGGCAATTAATCAAGCTCATCTGTAACGCTGTGAACGGGGGCTTTTCATCAAT 619
Db 121 LeuLysAlaTyrAsnGlnThrIleLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTCATTTGAATTTGACATCATCTGAGAGACAATATTTTAACTGGAGAACTCA 679
Db 141 CysThrTyrIleGlnIleGlyHisIleHisProGlyAspAsnIlePheLysLeuGlnSer 160
QY 680 CATTTTGAAGACGCGCGTGGAGAGATCCATATGACCTTAAGCTGTAACATCCCTT 739
Db 161 HisPheGlyAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrIlaSerLeu 180
QY 740 TTAATAGATGAGATTTATCTCTGAGACTGACGTGATTTTATGCGCGAGACTTCTCT 799
Db 181 LeuIleAspGlyLysLeuTyrSerIleTyrAlaIlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCGAAGACTCTTGGGACACCAACCAATCAGAGACAGACATGATTCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisIleHisProIleArgThrGlnGlnHisAspSerArgTrp 220
QY 860 CTCATATGATCCAAAGTTATTAATGTCCTCATCTCAGAGAGTGAATCTTGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerIlaHisLeuIleSerGlySerAspAsnProGlyAsp 240
QY 920 GACAAAGTATCTTTTCTCCGTAAGAAATGAATGAGATGAGAACCTCTGGAAGACT 979
Db 241 AspLysValTyrPhePhePheArgGlnAsnAlaIleAspGlyGlnHisSerGlyLysAla 260
QY 980 ACTCAGCTAGAAATAGTCAATATGCAAGATGACCTTGGAGGGCAGACAGAGTCTGGT 1039
Db 261 ThrHisAlaArgIleGlyGlnIleLysValAsnAspPheGlyGlyHisArgSerIleVal 280
QY 1040 AATAAATGACAACTTCTCAAGCTGCTGATTTGCTCAGTCCAGTCCAAATGTC 1099
Db 281 AsnLysTyrPheThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACATCTATTTTGAAGACTGAGAGATGATATTCCTTAAGAACTTAAAGATCTTAA 1159
Db 301 IleAspThrHisPheAspGlyLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCAGTTGTAATAGAGTGTATTACGATTCAGATTAATTTTCAAGAGATACGCCGTG 1219
Db 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerIlaVal 340
QY 1220 TGTATGATATGATGATGATGATGAGAAAGGATGTTCTGTGTCATATGATCCACAGGAT 1279
Db 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisIleArgAsp 360

```


TITLE OF INVENTION: Semaphorins
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSEE: RICHARD ARON OSMAN, Ph.D.
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/488,935
FILING DATE:
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/489,057
FILING DATE: 09-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-2/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 771 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-488-935-2
Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: Gaps: 0
US-09-774-490-1 (1-2709) x US-08-488-935-2 (1-771)
QY 200 ATGGCTGTGTTAATGAGTGTCTGCTTTCTGGGAGTATTACTTACAGCAAGACA 259
DB 1 MetGlyTTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTTCAGAAATGGGAAGAAACAATGTCAGAGCTGAATTTATCCACAAAGAAATGTG 319
DB 21 AenIYrgInAsnGlyLysAsnAsnValProArgLeuValLeuSerTyrIysGlnWetLeu 40
QY 320 GAATCAACAATGTGATCTTCAATGAGCTTGCCACAGCTCCAGTTATCATACCTTC 379
DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyrHisThrPhe 60
QY 380 CTTTGGATGAGAAACGAGTGGCTGTATGTTGAGCAAGATCATATTTTCATTC 439
DB 61 LeuLeuAspGlnGluLysSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCGGTAAATATCAAGATTTTCAAAAGATTTGTGCGCAGTATCTTACACCAAGACA 499
DB 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrrProValSerTyrThrArgArg 100
QY 500 GATGATGCAAGTGGCTGGAAGAAAGCATCCGGAAGAAATGCTATTTATTCACAGGTA 559
DB 101 AspGlnGlySerTrpAlaGlyLysAspIleLeuLysGlnCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGCATATATATCAAGCTCATTTGACGCTGTGAAACGGGGCTTTTCATCCAAAT 619

DB 121 LeuValAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTCATTTGAATTTGACATCATCTCGAGCAATATTTTAACTGAGAACCTCA 679
DB 141 CysThrTyrIleGlnIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTCGAAAACGGCGTGGGAAGATCATATGACCTTGAAGCTGTCAGACATCCCTT 739
DB 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATAGATGAGAAATTAATCTGTGAACCTGAGTATTTTATGGCGGAGACTTTGCT 799
DB 181 LeuIleAspGlyGlnLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTCCGAACTCTGGGACCAACCACTGAGCAAGACACATGATTCACAGTGG 859
DB 201 IlePheAsnGlnLeuGlyHisHisHisProIleArgGlnGlnHisAspSerArgTrp 220
QY 860 CTCAATGATCCAAAGTCAATATGATGCGCACTCATCTGAGAGTGAACATCTGAAAGAT 919
DB 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTAACTTTTCTCCGTGAAGAAATGCAATGAGAAACCTCTGGAAGAGCT 979
DB 241 AspLysValTyrPhePhePheArgGlyLysAsnAlaIleAspGlyGlnHisSerGlyLysAla 260
QY 980 ACTCAAGTAAATGTCAGATATGCAAGAAATGACTTTGAGGGCACAGAGTGTGCTG 1039
DB 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAAATGAGCAACATTTCTCAAACTGCTGATTTGCTCAGTCCAGTCCAAATGAGC 1099
DB 281 AenLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTATTTGATGAACTGCAAGAGTATTTCTTAATGAACTTTAAATCCCTAA 1159
DB 301 IleAspThrHisSphaAspGlnLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCGAGTGTATATGAGTGTGTTACGATTCCTCAATACATTTTGAAGGATCACCGCG 1219
DB 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATGACATAGATGATGAGAAAGGATGCTTGTGCTCATATGACCACAGGAT 1279
DB 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCAACTATCAATGGGTGCTTATCAAGGAAGATCCCTATTCACGGCCAGGAAT 1339
DB 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCAGCAAAACATTTGTGTGCTTTTGAATCTTAAAGCACTTCCTGAGATGATTA 1399
DB 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGAAGAAGTATCCAGCTCAGCTATCAATCCAGTGTTCATGAAACAATGCCCA 1459
DB 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
QY 1460 ATAGTATCAAAACGAGTGAATTAATTAATTAACAATTTGTAAGACCGAGTGAT 1519
DB 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValIleAspArgValAsp 440
QY 1520 GGAGAAGATGACATATGATGCTTATGTTATCGGAAGAAGATGGGACGCTTTTAA 1579
DB 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTCAATTTCTCAAGAGACTGTGATGATTTTGAAGAGGTTCTGCTGGAAGAAATG 1639
DB 461 ValValSerIleProLysGlnThrTrrTyrAspLeuGlnGlnValLeuLeuGlnGlnMet 480
QY 1640 ACAGTTTTTCGGGAACGAGCTGTATTTTCAAGATGAGGCTTTTCACTAAGCAGCAAA 1699
DB 481 ThrValPheArgGluProThrAlaIleSerAlaMetGlnLeuSerThrLysGlnGlnGln 500

QY 1700 CTATATATGTTCAACGGCTGGGCTTGCCAGCTCCCTTACACCGGTGATATATAC 1759
| | | | |
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
| | | | |
QY 1760 GGGAAAGCGTGTCTGAGTGTGTGCTGCGCCGAGACCTTACTGTGCTGGATGTCTT 1819
| | | | |
Db 521 GlyValAlaCysAlaGlnCysCysLeuAlaArgAspProTyrCysAlaTPrAspGlySer 540
| | | | |
QY 1820 GCATGTTCTGGTATTTTCCCATCTGCAAGAAGACGCAAGAGAGACAAAGTATTAAGAAAT 1879
| | | | |
Db 541 AlaCysSerArgTyrPheProThrAlaValArgArgThrArgGlnArgGlnArgAsn 560
| | | | |
QY 1880 GGAGACCCACTGACTCTGCTGCAAGCTTACACCATGATATATACCATGCGACAGCCCT 1939
| | | | |
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580
| | | | |
QY 1940 GAAGAGAAATCATCTATGGGTGTAAGAGATATAGACATTTTGGATCCAGTCCGAG 1999
| | | | |
Db 581 GlnGlnArgIleIleTyrGlyValGlnAsnSerSerThrPheLeuGlnCysSerProLys 600
| | | | |
QY 2000 TCGCAGAGAGCGGTGTCTTGTGGCAATTCAGAGCGGAATGAAGCGCAAGAGAG 2059
| | | | |
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgArgGlnArgGlnArgGln 620
| | | | |
QY 2060 ATCAGAGTGGATGATCATATCATCAGACAGATCAAGGCTTGTGCTAGCTAGTCTACA 2119
| | | | |
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnIleuLeuLeuAspSerLeuGln 640
| | | | |
QY 2120 CAGAGAGATTCAGGCAATTCCTCTGCGCATGCGGTGAGACATGCGTTTCAACAACTTT 2179
| | | | |
Db 641 GlnLysAspSerGlyAsnTyrLeuGlyHisAlaValGlnIleGlyPheIleGlnThrLeu 660
| | | | |
QY 2180 CTTAAGATTAACCTTGAAGTCACTGACAGACATTTGGAAGAACTTCTCTTAAGAT 2239
| | | | |
Db 661 LeuLysValThrLeuGlnValIleAspThrGlnHisLeuGlnIleuLeuHisLysAsp 680
| | | | |
QY 2240 GATGATGAGATGGCTCTAAGACCAAGAAATGTCATAGCATGACACCTTACCCAGAG 2299
| | | | |
Db 681 AspAspGlyAspIleTyrLeuSerThrLysGlnMetSerAsnSerMetThrProSerGlnLys 700
| | | | |
QY 2300 GTCTGTCACAGACATTCATGACAGCTTCATCAACCAACCAATTCACACAGATGATGAG 2359
| | | | |
Db 701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGln 720
| | | | |
QY 2360 TTCTGTGAACAAGTTTGAAAGGACCGAAACAACGTCGGAAAGCGCAGACATACC 2419
| | | | |
Db 721 PheCysGlnGlnValIleTyrLysArgAspArgLysGlnArgGlnArgProGlyHisIleThr 740
| | | | |
QY 2420 CCAGGAGACATTAACAATGSAAGACCTTACAAGAAATTAAGAAAGTGAAGACAGAGG 2479
| | | | |
Db 741 ProGlyAsnSerAsnLysIleTyrHisIleLeuGlnGlnLysAsnLysLysGlyArgAsnArg 760
| | | | |
QY 2480 ACCCAGCAATTTGAGAGGCGCACCCAGAGAGTTC 2512
| | | | |
Db 761 ThrIleGlnPheGlnArgAlaProArgSerVal 771
| | | | |
RESULT 4
US-08-489-057-2
Sequence 2, Application US/08489057
GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Matthes, David
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: Semaphorins
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: RICHARD ARON OSMAN, Ph.D.
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA

COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/489,057
FILING DATE: 09-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-2/BAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 771 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-489-057-2
Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
Gaps: 0
US-09-774-490-1 (1-2709) x US-08-489-057-2 (1-771)
QY 200 ATGGCTGGTTTAACATAGAGATTGTCTGTTTCTGGGAGTATTAATTACAGCAAGAGA 259
| | | | |
Db 1 MetGlyTrpLeuThrThrArgIleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
| | | | |
QY 260 AACTATCAGATGGAAGAGACATATGTCACAGGCTGAATTAATCTTACAAAGAAATGTTG 319
| | | | |
Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrIleGlnMetLeu 40
| | | | |
QY 320 GAATCAACAATGATATCACTTTCAATGCTTGCCACACAGCTCCAGTATCATACCTTC 379
| | | | |
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyrHisThrPhe 60
| | | | |
QY 380 CTTTGGATGAGAGACGAGTAGAGCTGTATGTTGGAGCAAGAGATCACATATTTTCATTC 439
| | | | |
Db 61 LeuLeuAspGlnGlnLysSerThrArgLeuTyrValGlyAlaLysAspHisIleLeuPheSerPhe 80
| | | | |
QY 440 GACTGTTTATATATCAAGGATTTTCAAAAAGATTTGTGGCCAGTATCTTACACAGAGA 499
| | | | |
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
| | | | |
QY 500 GATGAATGAATGGCTGCGAAAAGACATCCCGAAGAAATGTCATATTTCAAGGTA 559
| | | | |
Db 101 AspGlnCysLysTrpAlaGlyLysAspIleLeuLysGlnCysAlaAsnPheIleLysVal 120
| | | | |
QY 560 CTTAAGCATATTAATCAGACTCATCTTACGCTGTGAGACGCGGGCTTTTCATCCAAAT 619
| | | | |
Db 121 LeuLysAlaIleTyrAsnGlnThrHisIleuLysTyrAlaCysGlyThrGlyAlaPheHisProIle 140
| | | | |
QY 620 TGCACCTTACATTAATGAATGAACATCATCATCTGAGCAATATTTTAAGCTGAGAACTCA 679
| | | | |
Db 141 CysThrTyrIleGlnIleGlyHisHisProGlnLysAsnIlePheLysLeuGlnLysSer 160
| | | | |
QY 680 CATTTTGAAGAGCGCGGAGAGAGTGCATATGACCTTAAGCTCTGACAGCATCCCTT 739
| | | | |
Db 161 HisPheGlnAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
| | | | |


```

QY 740 TTAATGATGAGAAATTATCTCTGAGACTGACGTGATTTATGAGGCGAGACTTGTCT 799
Db 181 LeuileaspGlyGluLeuLeuYserGlyThrAlaAlaaspPheMetCylArgaspPheAla 200
QY 800 ATCTCCGAACTCTTGGGACCAACCAATCAGACAGACAGATGATATTCAGGTGG 859
Db 201 IlePheArgThrLeuGlyAlaShiShiAspProIleArgThrGluGlnIleAspSerArgTyr 220
QY 860 CTCGAATGATCCAAAGTTCATGATGCCCCACCTCATCTCAGAGAGTGAATCTGAAGAT 919
Db 221 LeuAsnAspProLeuPheIleSerAlaShiSleuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTACTCTTTCTTCGCGTGAATAATGATGAGAACTCTTGAAAAGCT 979
Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTGAGAAATGCTGATGATGCAAGAAATGACTTTGGAGGGGACAGAACTGTGG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAATGAGACAAACATTCCTCAAGCTGCTGATGATTTGCTCAGTGCAGAGTCCAAATGSC 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATGCACTCATTTTATGATGAACTGCAAGATGATTCCTAATGAACTTTAAAGATCCTAA 1159
Db 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetCAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGTATGAGAGTGTTCACGATTCAGTACATTTTCAAGGGATCAGCCGTG 1219
Db 321 AsnProValIleValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTTAGCATGATGATGATGAGAGGGGTGCTTGTGTCATATGCCACAGAGAT 1279
Db 341 CysMetTyrSerMetSerAspValAlaArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCAACTATCAATGGGTGCTTATCAAGAGAGAGTCCCTATCCAGCGCAGAGAACT 1339
Db 361 GlyProAsnTyrGlnTyrValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCCAAGCAAAACATTGGTGGTTTGACTCTCAAAAGAGACCTTCTGATGATGTTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGGAGAGAGTCAATCCAGCCATGATCCAGTGTTCCTATGAAAGATGGCCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetCylArgAsnProValPheProMetCAsnAsnArgPro 420
QY 1460 ATAGTATCAAAACCGATGTAAATTATCAATTATCAAAATTTGCTAGACCGAGTGAT 1519
Db 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAAGTGCAGATGATGATGTTATGTTATCGAAACAGATGTTGGACCGTCTTAAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAAGTTCAATTCCTAAGAGAGCTGTATGATGTTAAGAGAGAGTCTGTGAAAGAAATG 1639
Db 461 ValValSerIleLeuLysGluThrTyrTyrAspLeuGluGlnValLeuLeuGlnGlnMet 480
QY 1640 ACAGTTTTTCCGGAACCGACTGCTATTTTCAGCAATGAGCTTCCACTAGAGCAACAA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
QY 1700 CTATATATTTGTTCAACGGCTGGGGTGGCCAGCTCCCTTNAACCGGTGTATATTTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GCGAAGAGGCTGTGAGTGTGCTGCGCGCGAGACCCCTTATGTGCTTGGATGCTTCT 1819
Db 521 GlyLysAlaCysAlaGluCysLysLeuAlaArgAspProTyrCysAlaIleTyrAspGlySer 540
QY 1820 GCATGTTCTCGCTATTTTCCCACTGCAAGAGACGCAACAAAGACGCAAGATATAAGAAAT 1879

```

```

Db 541 AlaCysSerArgTyrPheProThrAlaLysArgAlaGlnArgGlnAspIleArgAsn 560
QY 1880 GGAGACCCACTGACCTCAGCTTTGACACTTACACATGATTAACACACATGCGCCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisShiAspAsnHisShiGlyHisSerPro 580
QY 1940 GAAGAGAAATCATCTATGCTGTAGAGAAATGATACACATTTTGGAAATGACGTCCGAAG 1999
Db 581 GluGluArgIleIleTyrGlyValGluLysAsnSerThrPheLeuGluCysSerProLys 600
QY 2000 TCGCAGAGAGCGCTGCTTATTTGGCAATTCCAGAGCCGAAATGAAAGAGGAAAGAG 2059
Db 601 SerGlnAlaLeuValTyrTrpGlnPheGlnArgAlaGlnGluGluArgLysGluGln 620
QY 2060 ATCAGAGTGAATGATCATCATCATGACAGACATGAAGCCCTTCTGCTAGTACTACA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAGAGATTCAGCAATTAACCTGCGCATGCGGTGGAACATGGGTTCATACAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysShiAlaValGluHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAACTCCGTAAGTCAATTGACAAGAGCATTTGAAAGAACTTCTCATTAAGAT 2239
Db 661 LeuLysValThrLeuGluValIleAspThrGlnHisSleuGluGluLeuLeuHisLysAsp 680
QY 2240 GATGATGAGATGCTCTTCAACCAAGAAATGTCCATAGCATGACACTTGCCAGAG 2299
Db 681 AspAspGlyAspGlySerLysThrLysGlnLeuSerAsnSerMetThrProSerGlnLys 700
QY 2300 GTCTGTACAGAGACTTCATGAGAGTCAACCAACCCCAATCTCAACAGATGATGAG 2359
Db 701 ValTyrTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGln 720
QY 2360 TTCTGTGAACAAGTTTGAAGAAAGGACCGCAAAACACACTCGGCAAAAGCCAGACATACC 2419
Db 721 PheCysGluGlnValTyrLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
QY 2420 CCAAGGAACAGTAACTAAGATGAGACATTTACAGAAATTAAGAAAGTGAAGACGAGAG 2479
Db 741 ProGlyAsnSerAsnLysTyrLysHisSleuGlnGluAsnLysGlyArgAsnArgArg 760
QY 2480 ACCCAGAAATTTGAGAGGACCCCGAGAGTGC 2512
Db 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 5
US-09-694-085-24
; Sequence 24, Application US/09694085
; GENERAL INFORMATION:
; APPLICANT: Behar, Oded
; APPLICANT: Woolf, Clifford J.
; TITLE OF INVENTION: NOVEL SEMAPHORIN DOMAINS
; FILE REFERENCE: 00786-441001
; CURRENT APPLICATION NUMBER: US/09/694,085
; CURRENT FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/164,056
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-694-085-24

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0

```

DB: 20 Gaps: 0
US-09-774-490-1 (1-2709) x US-09-694-085-24 (1-771)
QY 200 ATGGGCTGGTTACTAGATGTTCTGCTTTCTGGGGAGTATTACTTACAGCAAGCA 259
DB 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpValLeuLeuThrAlaArgAla 20
QY 260 AACATCAGAAATGGAGAAACAATGTCGCAAGGCTGAAAATTATCTTACAAAGAAATGTTG 319
DB 21 AsnTyrGlnAsnGlyAsnAsnValProArgLeuLysLeuSerTyrLysGlnMetLeu 40
QY 320 GAATCCAAACATGTGATCATCTTCAATGCGTTGGCCAAACAGCTCCAGTTATCATACCTTC 379
DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyrHisThrPhe 60
QY 380 CTTTGGATGAGAAAGGAGTAGGCTGTATGTTGGAGCAAAAGATCATATTTTCACTTC 439
DB 61 LeuLeuAspGlyGlnLysSerArgLysValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAATATCAAGATTTTCAAAAGATTGTGGCCAGATCTTACACAGAAAG 499
DB 81 AspLeuValAsnIleLysAspPheGlnLysIleValIleProValSerTyrThrAlaArg 100
QY 500 GATGAATGCAAGTGGGCTGGAAAAGACATCTGAAAAGATGTCATATTTCAATCAAGTA 559
DB 101 AspGlyCysLeuTrpAlaGlyLysAspIleLeuLysGlyCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATATCAAGCTCCTGTACGCTGTGGAAAGCGGGGCTTTTCAATCCAAAT 619
DB 121 LeuLysAlaLysAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTACATTAATTAATGGACATCATCTGAGAGCAAAATTTTAACTGGAGAACTGA 679
DB 141 CysThrTyrIleGluIleGlyHisIleProGlnAspAsnIlePheLysLeuGlnAsnSer 160
QY 680 CATTTTGAACCGCGCTGGGAAGAGTCCATATGACCTTAAGCTGTGACAGCATCCCT 739
DB 161 HisPheGlyAsnGlyArgGlyLysSerProTyrAspProLysLeuThrAlaSerLeu 180
QY 740 TTAATGATGAGAAATTAATTAATCTGGAATGCGAGCTGATTTATGGGGCAGACTTGGCT 799
DB 181 LeuIleAspGlyGlyLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATCTTCGAACTCTTGGGCGACGACCAACCATGAGAGAGAGAGATGTTCCAGATGG 859
DB 201 IlePheArgThrLeuGlnHisIleHisProIleArgThrGlnGlnHisAspSerArgTrp 220
QY 860 CTCATGATCCAAAGTTTCAATAGTCCACCTCATCTCAGAGAGTGAACAATCCTGAAGAT 919
DB 221 LeuAsnAspProLysPheIleSerAlaHisIleuIleSerGluSerAsnProGlnAsp 240
QY 920 GACAAAGTATATCTTTTCTCCGTGAATGCAATGATGAGAGAAACCTCTGGAAGAGCT 979
DB 241 AspLysValTyrPhePhePheArgGlnAsnAlaIleAspGlyGlnHisSerGlyLysAla 260
QY 980 ATCCACGCTAGATAGATGATGCAAGATGACCTTGGAGGGGACAGAGAGTCTGGTG 1039
DB 261 ThrHisAlaArgGlyIleGlyGlnIleCysLysAsnAspPheGlyGlnHisArgSerLeuVal 280
QY 1040 AATAAATGACAACATCTCTCAAGCTCGTGTGATTTGCTCAGTGGCAGAGTCCAAATGCG 1099
DB 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACATCTATTTGATGAATGCGAGATGTAATTTCTTAATGAACCTTAAAGTCTTAA 1159
DB 301 IleAspThrHisPheAspGlyLeuGlnAspValPheLeuMetCAsnPheLysAspProLys 320
QY 1160 AATCAGTTGATATGAGAGTGTTTAGCATCCAGTAACATTTTCAAGGGATCAGCCGG 1219
DB 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGCATGATGATGTGAGAAAGGTGTTCTTGGTCCATATGCCACAGGAGAT 1279

DB 341 CysMetTyrSerMetSerAspValAlaArgValPheLeuGlyProTyrAlaHisArgAsp 360
QY 1280 GGACCCAACTATCATATGGGTGCTTATGCAAGAAAGATCCCTATCCAGGCGCAGAACT 1339
DB 361 GlyProAsnTyrGlnTrpAlaProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCAGCAAAACATTTGGTGGTTTGACTTACAAAGACCTTCTGATATGTTATA 1399
DB 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGAAGAGTCAATCCAGCCATGTAACAATCCAGTGTTCCTATGAAACATCGCCA 1459
DB 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetCAsnAsnArgPro 420
QY 1460 ATATGATCAAAACGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1519
DB 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAAGATGACAGATGATGTATGTTATCGGAACAGATGTTGGACCGTCTTAA 1579
DB 441 AlaGlyAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTATGTTCAATTCCTAAGAGACCTTGATGATTAAGAAAGGTTCTGCGAAGAAATG 1639
DB 461 ValValSerIleProLysGlyThrTrpTyrAspLeuGlnGlnValLeuLeuGlnGlnMet 480
QY 1640 ACGATTTTCCGGAACCGCATCTGCTTTCAGAAATGAGACCTTCCACTAAGAGAGCA 1699
DB 481 ThrValPheArgGlyProThrAlaIleSerAlaMetGlnLeuSerThrLysGlnGlnGln 500
QY 1700 CTATATATGTTCAACGGCTGGGGTGGCCAGCTCCCTTACACCGGATGATATTAC 1759
DB 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY 1760 GGGAAAGCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTTGGATGATCT 1819
DB 521 GlyLysAlaCysAlaGlyCysCysLeuAlaArgAspProTyrCysAlaIleTrpAspGlySer 540
QY 1820 GCATGTTCCGCTATTTTCCCATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1879
DB 541 AlaCysSerArgTyrPheProThrAlaLysAspArgThrAlaArgGlnAspIleArgAsn 560
QY 1880 GGAGACCACTGATCACTGCTTCAAGCTTACACATGATTAATCACATGAGCCAGACCT 1939
DB 561 GlyAspProLeuThrHisCysSerAspLeuHisIleAspAsnHisIleGlyHisSerPro 580
QY 1940 GAAAGAGATCATGTATGTTGATAGAAATAGTACCAATTTTGAATGCAAGTCCGAAG 1999
DB 581 GlnGlnArgIleIleTyrGlyValGlnAsnSerSerThrPheLeuGlnCysSerProLys 600
QY 2000 TGGCAGAGAGCGCTGGTCTATTTGGCAATTCAGAGAGCGAAATGAGAGCGAAAGAG 2059
DB 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnAspArgAspGlnAlaArgLysGlnGln 620
QY 2060 ATCAGAGTGAATGATCATATCATCAGACAGATCAAGGCTTCTGCTAGTGTACAA 2119
DB 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuHisArgSerLeuGln 640
QY 2120 CAGAGAGATTCAGGCAATTAATCTGCGATGGGTGGAACATGGTTCATACAAATCTTT 2179
DB 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGlnHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGTAAACCTTGAAGTCAATGACACAGAGCTTTGGAAGAACTTTCATTAAGAT 2239
DB 661 LeuLysValThrLeuGlnValIleAspThrGlnHisLeuGlnGlnLeuLeuHisLysAsp 680
QY 2240 GATGATGAGATGGCTCTAAGACCAAGAAATGTCATATGATGATGACACTTACAGCAAG 2299
DB 681 AspAspGlyAspGlySerLysThrLysGlnMetSerAsnMetThrProSerGlnLys 700
QY 2300 GTCTGTACAGAGACTTCAATGAGCTCATCAACCAACCAATTCACAGATGATGAG 2359

Db 701 ValTrpTyArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrIleAspGlu 720
Qy 2360 TTCTGTGAACAAGTTTGGAAAAGGACCGGAAAACAAGTGGCAAGGCGACAGACATACC 2419
Db 721 PheCysGlnGlnValTrpTyArgAspArgLysGlnArgGlnArgGlnArgProGlnHisThr 740
Qy 2420 CCAAGGACAGATGACAAATGGAAGACCTTACAGAAAATAGAAAGTGTGAAAACAGAGG 2479
Db 741 ProGlnAsnSerAsnLysTrpLysHisIleGlnGlnLeuAsnLysGlnArgAsnArgArg 760
Qy 2480 ACCCAGCAATTTGAGAGGCGCACCCAGAGTGTG 2512
Db 761 ThrHisGlnPheGlnArgAlaProArgSerVal 771
RESULT 6
US-09-791-537-31106
; Sequence 31106, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 31106
; LENGTH: 771
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-791-537-31106
Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 22 Gaps: 0
US-09-774-490-1 (1-2709) x US-09-791-537-31106 (1-771)
Qy 200 ATGGGCGTGTACTAGAGATGTCGTCTTTCTGGGAGATTAATCTTACAGAAAGAGA 259
Db 1 MetGlyTrpLeuThrArgLleValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20
Qy 260 AACTATCAGAAATGGAAGAACAATGCGCAAGGCTGAATTAATCTTACAAAGAAATGTTG 319
Db 21 AsnTrpGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTrpLysGlnMetLeu 40
Qy 320 GAATCCAAATATGTGATCACTTTCAATGGCTTGGCCAAAGCTCCAGTTATCAATACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTrpHisThrPhe 60
Qy 380 CTTTGGATGAGGAACGAGTAGTGGCTGTATGTTGAGCAAGAAATACATATTTTCATTC 439
Db 61 LeuLeuAspGlnGlnArgSerArgLeuTrpValGlyAlaLysAspHisIlePheSerPhe 80
Qy 440 GACGTGTTAATATCAAGATTTTCAAAAGATTTGTGTGCGCAGTATCTTACACAGAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTrpThrArgArg 100
Qy 500 GATGAATGCAAGTGGCTGGAAGAAAGACATCTCGAAGAAATGTGCTAATTTTCATCAAGTA 559
Db 101 AspGlnCysValSerTrpAlaGlyLysAspIleLeuLysGlnCysAlaAsnPheIleLysVal 120
Qy 560 CTTAAGGCAATATATAGACTCACTGTATGCGCTGTGGAACGGGGCTTTTCATCCAAAT 619
Db 121 LeuLysAlaTrpAsnGlnThrHisLeuTrpAlaCysGlyThrGlyAlaPheHisProIle 140
Qy 620 TCACCTACATTCGAATTTGACATCATCTCTGAGAGCAATATTTTAACTGAGAACTCA 679

Db 141 CysThrTrpLleGlnIleGlyHisHisProGlnAspAsnIlePheLysLeuGlnLysAsnSer 160
Qy 680 CATTTTGAACCGGCGCGGGAAGAGTCATATGACCTTAAGCTCTGACAGACATCCCTT 739
Db 161 HisPheGlnAsnGlyArgGlyLysSerProTrpAspProLysLeuLeuThrAlaSerLeu 180
Qy 740 TTAATAGATGAGATTTATATCTGTGAACCTGAGCTGATTTTATGAGGCGAGACTTTGCT 799
Db 181 LeuIleAspGlyGlnLeuTrpSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
Qy 800 ATCTTCGAACCTTTGGGACACCCCAATCAGACAGACAGACAGATTCAGGTGG 859
Db 201 IlePheArgThrLeuGlyHisHisHisProIleArgTrpGlnGlnHisAspSerArgTrp 220
Qy 860 CTCATATGATCCAAATTCATTAATGTCGCCACCTCATCTCAGAGACATGACATTCGAAGT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGlnSerAspAsnProGlnLysP 240
Qy 920 GACAAAGTATACTTTTCTCCGTGAAGAAATGCAATAGATGAGAGACACTGTGAAAAGCT 979
Db 241 AspLysValTrpPhePhePheArgGlnAsnAlaIleAspGlyGlnHisSerGlyLysAla 260
Qy 980 ACTCAGCTAGAAATGTCAGATATGCAAGAAATGACTTTGAGGGCAGAGAGTCTGTG 1039
Db 261 ThrHisAlaArgLleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
Qy 1040 AATAATAGACAACTTCTCAGAACTGCTGATTTGCTCAGTCCAGTCCAAATGTC 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
Qy 1100 ATTGACACTCAATTTGATGAACGTGACAGATGATTCCTAATGAATTTAAGATCCTTAA 1159
Db 301 IleAspThrHisAspAspGlnLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
Qy 1160 AATCCAGTTGATATGAGAGTGTTAACAATTCAGATTAATTTTCAAGAGATCAGCCGTG 1219
Db 321 AsnProValValTrpGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
Qy 1220 TGTATGTAATGATAGATGATGATGAGAGGAGTTCCTTGGTCCATATGCCACAGGAT 1279
Db 341 CysMetTrpSerMetSerAspValArgArgValPheLeuGlyProTrpAlaHisArgAsp 360
Qy 1280 GGACCCCACTAATCAATGGGTGCTTATGACAGAAAGTCCCTATCCACGGCAGAGAACT 1339
Db 361 GlyProAsnTrpGlnTrpValProTrpGlnGlyArgValProTrpProArgProGlyThr 380
Qy 1340 TGTCCAGCAAAACATTTGGTGTGTTGACTCTACAAAGACCTTCAGATGATGTTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
Qy 1400 ACCTTTGAAGAAGTCATCCAGCCCATGTAACAATCCAGTGTTCCTATGAAACATGCCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTrpAsnProValPheProMetAsnAsnArgPro 420
Qy 1460 ATAGATGCAAAAGCGATGATAATATGATTAACAATTTGAGTGTGACCGAGTGGAT 1519
Db 421 IleValIleLysThrAspValAsnTrpGlnPheThrGlnIleValValAspArgValAsp 440
Qy 1520 GCAGAGATGACAGATATGATGTTATGTTATCGCAACAGATGTTGGACCGTCTTAA 1579
Db 441 AlaGlnAspGlyGlnTrpAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
Qy 1580 GTAATTTCAATTCCTAAGAGACTTGGATGATTTTGAAGAGGTTCTGCTGGAAGAAATG 1639
Db 461 ValValSerIleProLysGlnThrTrpTrpAspLeuGlnGlnValLeuLeuGlnLysMet 480
Qy 1640 ACAGTTTTCGGAACCGACTGCTATTTCAAGATGAGGCTTTCACATAGAGCAACA 1699
Db 481 ThrValPheArgGlnProThrAlaIleSerAlaMetGlnLeuSerThrLysGlnGlnGln 500
Qy 1700 CTATATATGGTTCAACGGCTGGGGTTGCCAGCTCCCTTTACACCGGTGTGATATTTAC 1759

```

Db      501 LeuTyrlleGlyserThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTy 520
QY      1760 GGAAGAGCGTGTGTGAGTGTGTGCTGCGCCGAGACCTTACTGTGTGGATGGTCTT 1819
Db      521 GilyValaIaCysalaGluCysCysLeuAlaArgAspProTyrcysAlaIaTpaapGlySer 540
QY      1820 GCATGTTCTGCTATTTTCCCACTGCAAGAAGACGCAAGACGACAAAGATATAGAAAT 1879
Db      541 AlAcYserArgTyrrhrProThrAlaIleYhrArgThrArgAlaGlnAlaIleArgAsn 560
QY      1880 GGAAGACCACTGACTACTGTTCAAGACTTACACCATGATATACCATCGCCACAGCCT 1939
Db      561 GilyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580
QY      1940 GAAGAGGATTCATCTTGTGTGTAGAGATAGTACATTTTGGATTCAGTCCGAG 1999
Db      581 GlnGlnArgIleIleTyrgIyValGlnAsnSerSerThrPheLeuGluCysSerProLys 600
QY      2000 TCCGAGAGACCGCTGTCTTATTTGGCAATTCAGAGCGGAATGAGAGCAAGAAAGAG 2059
Db      601 SerGlnArgAlaLeuValTyrrTpgInPheGlnArgArgAsnGlnGlnArgIySGlnGln 620
QY      2060 ATCAGAGTGAATGATCATCATCATCAGACAGATCAGAGCCTTCTGCTTACTTCAAT 2119
Db      621 IleArgValaAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY      2120 CAGAAGATTCAGAGCAATTAACCTCTGCGCATGCGGTGAGACATGGGTTTCATCAAC 2179
Db      641 GlnLysAspSerGlyAsnTyrcLeuCyHisAlaValAlaIleIleGlyPheIleGlnThrLeu 660
QY      2180 CTTAAGGTTAACCTGTGAGAGTCAATTCAGACAGACATTTGGAGAGAACTTCTTCA 2239
Db      661 LeuIyValIleThrLeuGlnValIleAspThrGlnHisLeuGlnGlnLeuLeuHisIyAsp 680
QY      2240 GATGATGAGATGAGCTTAAAGACCAAGAAATGTCATATGACATGACACTTACCGAGAG 2299
Db      681 AspAspGlyAspGlySerIySthrIySGlnMetSerAsnSerMetThrProSerGlnLys 700
QY      2300 GTCTGTACAGACATTCATGACAGCTCATCAACACCCCAATTCACAGATGAGTGA 2359
Db      701 ValTyrrTyrrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGln 720
QY      2360 TTCGTGACCAAGTTTGGAAAAGGACCGAAAACAACGTGCGCAAAAGCCAGACATATAC 2419
Db      721 PheCysGlnGlnIleValIleTyrrAspArgAspArgIySGlnArgArgGlnArgProGlySthr 740
QY      2420 CCGAGGAACGATTAACAATGGAAGCACTTACAAAGAAATTAAGAAAGGTAGAAAACAGAG 2479
Db      741 ProGlyAsnSerAsnIySthrIySleuGlnGlnIleAsnIySGlyArgAsnArgArg 760
QY      2480 ACCCAGCAATTGAGAGGGGACCCAGAGAGTGC 2512
Db      761 ThrHisGlnPheGlnArgAlaProArgSerVal 771

RESULT 7
US-10-067-632-54
; Sequence 54, Application US/10067632
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
;               Kolodkin, Alex L.
;               Matthes, David
;               Bentley, David R.
;               O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSER: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/067,632
; FILING DATE: 04-Feb-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,610
; FILING DATE: <Unknown>
; APPLICATION NUMBER: 08/835,268
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 771 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 54:
US-10-067-632-54

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 26 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-067-632-54 (1-771)
QY      200 ATGGCGTGGTTAATGAGATTTGTCTGTTTGGGAGATTAATCAAGCAAGACA 259
Db      1 MetGlyTyrrLeuThrArgIleValCysLeuPheThrGlyValLeuLeuThrAlaArgAla 20
QY      260 AACTCTGCAAGAGGAGAAACAATGTCGCAAGCTGGAATTAATCTTACAAAGAAATGTTG 319
Db      21 AsnTyrgIleAsnGlyLysAsnAsnValProArgLeuIySleuSerTyrrSGlnMetLeu 40
QY      320 GAATCCAACAATGTATGATCACTTCAATGAGCTTGCCACAGCTCCAGTTATCATACCTTC 379
Db      41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrrHisThrPhe 60
QY      380 CTTTGGATGAGAAACGAGTARGCTGTATGTTGGAGCAAAAGATCAATATTTCAATTC 439
Db      61 LeuLeuAspGlnGlnArgSerArgLeuTyrrValGlyAlaIyAspHisIlePheSerPhe 80
QY      440 GACCTGTTAATATATCAAGATTTTCAAAAAGTTGTGCGCAAGTATCTTACACCGAGAGA 499
Db      81 AspLeuValaAsnIleLysAspPheGlnIySleuIleValIleProValSerTyrrThrArgArg 100
QY      500 GATGATGACAGTGGGCTGGAAGAAAGACATCCGAAAGATGTCTTAATTTTCATCAAGTA 559
Db      101 AspGlnCysIySthrAlaGlyLysAspIleLeuIySGlyCysAlaAsnPheIleIySthrVal 120
QY      560 CTTAAGCATATATATCAACTGCTTGAACGCTGTGGAACGGGGGCTTTTCATCAATTT 619
Db      121 LeuIySalIyArgAsnGlnThrHisIleuTyrrAlaCysGlyThrGlyAlaPheHisProIle 140
QY      620 TGCACCTCATTTGAATGAGACATCATCTGAGAGCAATATTTTAACTGGAGAACTCA 679
Db      141 CysThrTyrrIleGlnIleGlyHisIleProGlnAspAsnIlePheIySleuGlnAsnSer 160
QY      680 CATTTGAAAACGGCGGTGGAAGAGTCCATATGACCTTAAGCTGTGACAGATCCCTT 739

```

```

Db      ||||| 161 HispGluAspGlyArgGlySerProTyrAspProLysLeuThrAlaSerLeu 180
Qy      ||||| 740 TTAATGATGAGAAATTAATCTGGAACCTGAGATTTTATGGGGGAGACTTGGT 799
Db      ||||| 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheMetIleYArgAspPheAla 200
Qy      ||||| 800 ATCTTCGAACCTTGGGACCAACCAATCAGAGACAGACGATATTCAGAGTGG 859
Db      ||||| 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTrp 220
Qy      ||||| 860 CTCATGATCCAAAGTTCAATGATGCCACCTTCATTCAGAGATGACAACTCTGAAGAT 919
Db      ||||| 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
Qy      ||||| 920 GACAAAGTATCTTTCTCCGTGAAATGCAATGATGAGGAAACCTCTGAAAGT 979
Db      ||||| 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
Qy      ||||| 980 ACTCAGCTAGAAATAGTACATATGCAAGATGACTTTGAGGGGACGAAGTCTGTG 1039
Db      ||||| 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
Qy      ||||| 1040 AATTAATGCAACAATCTCTCAAGCTCGTGTGATTTGCTCAGTGCAGGTCCAAATGCG 1099
Db      ||||| 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAlaGly 300
Qy      ||||| 1100 ATTGACACCTATTTGATGAACTGACGAGATGTATTCCTTAATGAACCTTAAAGTCTTAA 1159
Db      ||||| 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
Qy      ||||| 1160 AATCAGGTGTATGATGAGTGTGTTACGACTTCAGTACATTTTCAAGGAGTCCAGCCGG 1219
Db      ||||| 321 AsnProValValTyrGlyValPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
Qy      ||||| 1220 TGTATGTATGATGATGATGATGATGAGAGGGTGTTCCTTGGTCCATATGCCACAGGAT 1279
Db      ||||| 341 CysMetTyrSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
Qy      ||||| 1280 GGAACCAACTATCATATGAGTGGTCTTATCAAGAGAGTCCCTTATCCAGGCCAGAGACT 1339
Db      ||||| 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
Qy      ||||| 1340 TGTCCAGCAAAACATTTGGTGTGTTGACTCTACAAAGACCTTCCATATGATTTATA 1399
Db      ||||| 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400
Qy      ||||| 1400 ACCTTGCAAGAGTATCAGCCAGCCATGTAACAATCCAGTGTTCCTATGAACAATGCCCA 1459
Db      ||||| 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
Qy      ||||| 1460 ATAGTATCAAAACGATGTAATTAATTCATTTACAAATTTGCTGAGACCGAGTGAAT 1519
Db      ||||| 421 IleValIleLysThrAspValAsnTyrGlnPheThrGlnIleValAlaAspArgValAsp 440
Qy      ||||| 1520 GCAGAGATGACAGATGATGTTATGTTATCGAAGACAGATTTGGGACCGTCTTAA 1579
Db      ||||| 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
Qy      ||||| 1580 GTAGTTTCATTTCTAAGAGACCTTGATGATTTAGAAAGAGTTCTGCTGAGAAAGATG 1639
Db      ||||| 461 ValValSerIleProLysGluThrTrpTyrAspLeuGluGluValLeuLeuGluGluMet 480
Qy      ||||| 1640 ACAATTTTGGGAAACCGACTGCTATTTAGCAATGAGACTTCCACTAGACGACAA 1699
Db      ||||| 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
Qy      ||||| 1700 CTATATATTGTTCAACGGCTGGGGTGGCCAGCTCCCTTCAACGGGTGATATTATTC 1759
Db      ||||| 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
Qy      ||||| 1760 GGGAAACGCTGTGCTGAGTGTTCCTCGCCGACGACCTTACTGTGCTGGAGTGTCT 1819

```

```

Db      ||||| 521 GlyLysAlaCysAlaGluCysLeuAlaArgAspProTyrCysAlaTrpAspGlySer 540
Qy      ||||| 1820 GCATGTTCCGTATTTTCCACTGCAAGAGAGCCACAAAGCAAGATTAAGAAAT 1879
Db      ||||| 541 AlaCysSerArgTyrPheProThrAlaLysArgGlnTrpArgGlnAspIleArgAsn 560
Qy      ||||| 1880 GGAGACCACTGACTCAGCTGTTCAGACTTAAACATGATTAATCACATGAGCCACAGCCT 1939
Db      ||||| 561 GlyAspProLeuThrHisCysSerThrPheLeuHisHisAspAsnHisGlyHisSerPro 580
Qy      ||||| 1940 GAGAGAGATCATTTATGCTGTAGAGAAATGACCAATTTTGAATTCAGTCCGAG 1999
Db      ||||| 581 GluGluArgIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
Qy      ||||| 2000 TCGCAGAGAGCGCTGCTATTTGCAATTCAGAGCGGAAATGAAGACGAAAGAG 2059
Db      ||||| 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgGlnArgGluArgGlyGluGlu 620
Qy      ||||| 2060 ATCAGAGTGAATGATCATATGATCATCAGACAGATCAAGGCTTCTGCTAGTGTACAA 2119
Db      ||||| 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
Qy      ||||| 2120 CAGAAAGATTAGGCAATTACTCTGCGATGCGGTGAACATGGGTTATCAAACTCTT 2179
Db      ||||| 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValGlnHisGlyPheIleGlnThrLeu 660
Qy      ||||| 2180 CTTAAGTATACCTCGGAAGTCAATGACACAGAGCTTTGGAAGAACTTCTCATTAAGAT 2239
Db      ||||| 661 LeuLysValThrLeuGlnValIleAspThrGlnHisLeuGlnGluLeuLeuHisLysAsp 680
Qy      ||||| 2240 GATGATGAGAGATGGCTCTAAGACCAAAAGATGTCCAATACATGACATGACCAAG 2299
Db      ||||| 681 AspAspGlyAspGlySerLysThrLysGlnMetSerAsnSerMetThrProSerGlnLys 700
Qy      ||||| 2300 GTCTGTACAGAGACTTCATGACGCTCATCAACCACTTCAACGATGATGATGATG 2359
Db      ||||| 701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
Qy      ||||| 2360 TTCGTGAAACAAGTTTGGAAAGGACCGAAACAAGCTCCGCAAGAGCCAGACATACC 2419
Db      ||||| 721 PheCysGluGlnValTyrLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
Qy      ||||| 2420 CCAGGGAACAGTAAACAATGAGACCTTACAGAAATTAAGAAAGATGAGAAACAGAGAG 2479
Db      ||||| 741 ProGluAsnSerAsnLysTrpLysHisLeuGlnGluAsnLysLysGlyArgAsnArgArg 760
Qy      ||||| 2480 ACCCAGATTTGAGAGGGCACCCAGAGTGTCT 2512
Db      ||||| 761 ThrHisGluPheGluArgAlaProArgSerVal 771

RESULT 8
US-10-097-340-284
; Sequence 284, Application US/10097340
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNANVARAPU
; APPLICANT: Sebastian HOERSCH
; APPLICANT: Shudhangit KAMATKAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIRY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030

```

```

; CURRENT APPLICATION NUMBER: US/10/097,340
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 284
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-097-340-284

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 26 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-097-340-284 (1-771)

QY 200 ATGGCGTGTACTAGAGATGCTGCTCTTTCTGGGAGATTAATTACAGCAAGCA 259
DB 1 MetGlyTTrpLeuThrArgIleValCysLeuPheTrrpGlyValLeuLeuThrAlaArgAla 20
QY 260 AACTACAGAAATGGAAGAACATATGACCAAGGCTGAAATTAATCTACAAAGAAATGTTG 319
DB 21 AenlyrGlnhensilylshasnasnvalProArgLeuLeuLeuSerTyrlYsgIuueLeu 40
QY 320 GAATCCAAAGATGATGATCACTTTCATGAGCTTGCCCAAGCTCCAGTTATCAATACCTTC 379
DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerSerTyrlHisThrPhe 60
QY 380 CTTTGGATGAGAGACGAGATGAGCTGTATGTTGGACCAAGATCAATATTTCAATTC 439
DB 61 LeuLeuAspGluGluArgSerArgLeuTyValGlyAlaIysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAATATCAAGAGATTTTCAAAAGATTTGTGGCCAGATCTTACACAGCAAG 499
DB 81 AspLeuValAsnIleLysAspPheGlnLysIleValIleProValSerTyrlHisArgAla 100
QY 500 GATGAATGCAAGTGGCTGGAAGAACATCTCGTGAAGAAATGTCATATTTCAAGGTA 559
DB 101 AspGluCysLeuSerTrpAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleLysVal 120
QY 560 CTTAAGGCATATATCAAGATCACTTGTACGCTGTGGAAACGGGGGCTTTTCATCCAAAT 619
DB 121 LeuLysAlaIleLysAsnGlnThrHisIleuTyValCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGACACTAGATTAATGAGATGACATCCCTGAGAGAAATTTTAAAGCTGAGAGACTGA 679
DB 141 CysThrTyrlIleGlnIleGlyHisIleProGlnAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTTGAAGACGCGCTGGAGAGAGATCCATATGACCTTAAGCTGTGACAGCATCCCTT 739
DB 161 HisPheGluAsnGlyArgGlyLysSerProTyLAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATATGATGAGAAATTAATCTTGGAATGCAAGCTGATTTTATGGGGGAGACTTTGCT 799
DB 181 LeuIleAspGlyGluLeuTyrlSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
```

```

QY 800 ATCTCCGAACCTTTGGGACCAACCACCAATACAGACAGACAGATGATTCAGAGTGG 859
DB 201 IlePheArgThrLeuGlyHisIleHisProIleArgGlnGlnHisIleAspSerArgTrp 220
QY 860 CTCATATGATCAAGTCAATTAATGTCGCCACCTCATCTCAGAGATGACATCTTAAGAT 919
DB 221 LeuAsnAspProLysPheIleSerAlaHisIleuIleSerGluSerAspAsnProGluAsp 240
QY 920 GACAAAGTATTAATTTTCTCCGTGAATAATGAAATGAGAACATCTTGGAAGAGT 979
DB 241 AspLysValTyrlPhePhePheArgGlnAsnAlaIleAspGlyGlnHisSerGlyLysAla 260
QY 980 ACTCAAGCTAGAAATGATGATGATGCAAGAAATGACTTTGGAGGCAAGAAATGCTGGT 1039
DB 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATTAATGACAACATTCCTCAAGCTGCTGATTTGCTCAGTCCAGATCAAGGATGCG 1099
DB 281 AenlystrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACACTCATTTTGAATGAATGACAGATGATTCCTTAATGAATTTTAAGATCTTAA 1159
DB 301 IleAspThrHisPheAspGluLeuGlnAspValIlePheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGTATGAGTGTGTTACGACTTCCAGATTAACATTTTCAAGGATCAAGCGTG 1219
DB 321 AsnProValValTyrlGlyValAlaPheThrTrpSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGATATGATGATGATGATGAAAGGCTGTTCTTGCTCAATATGCCCAAGGAT 1279
DB 341 CysMetTyrlSerMetSerAspValArgArgValIlePheLeuGlyProTyrlAlaHisArgAsp 360
QY 1280 GGAACCAATCATATGAGTGTGCTTATCAAGAAAGTCCCTATCCAGCGCAGAACT 1339
DB 361 GlyProAsnTyrlGlnTrpValProTyrlGlnGlyArgValProTyrlProArgProGlyThr 380
QY 1340 TGTCCAGCAAAACATTTGGTGTGTTTGAATCTTCAAAAGAACCTTCGATGATGTTATA 1399
DB 381 CysProSerLysThrPheGlyGlyPheAspSerThrTyLAspLeuProAspAspValIle 400
QY 1400 ACCTTGCAAGAGATCATCCACCATGTAACAATCAAGTGTTCCTATGAACATGCCCA 1459
DB 401 ThrPheAlaArgSerHisProAlaMetTyrlAsnProValIlePheProMetAsnAsnArgPro 420
QY 1460 ATAGTATCAAAACGGATGTAATTAATTAACAATTTGATGAGACCGAGATGAT 1519
DB 421 IleValIleLysThrAspValAsnTyrlGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAAGATGACAGTATGATGTTATGTTATCGGAACAGATGTTGGACGCTTTTAA 1579
DB 441 AlaGluAspGlyGlnTyrlAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCATATCTTAAGAGAACTTGATGATTTTGAAGAGGTTCTGCTGAGAGAAATG 1639
DB 461 ValIleSerIleProLysGluThrTyrlTyrlAspLeuGlnGluValLeuLeuGlnGluMet 480
QY 1640 ACAATTTTTCGGGAACCGATGCTATTTTCAGCAATGAGAGCTTTCCATTAAGAGAAAGA 1699
DB 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
QY 1700 CTATATATTTGGTTCAACGGCTGGGGGTTGCCAGCTCCCTTTTACACCGGCTGATTTTAC 1759
DB 501 LeuTyrlIleGlySerThrAlaGlyAlaAlaGlnLeuProLeuHisArgCysAspIleTyrl 520
QY 1760 GGGAAAGCGTGTGCTAGAGTGTGCTGCGCCGAGACCTTAATCTGTGCTGGAGTGTCT 1819
DB 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrlCysAlaIleTrpAspIleYer 540
QY 1820 GATGCTTCTCGTATTTTCCACTGCAAGAAAGACCAACAAGCAACAGATTAAGAAAT 1879
DB 541 AlaCysSerArgTyrlPheProThrAlaLysArgArgTrpArgArgGlnAspIleArgAsn 560
```



```

QY 1220 TGTATGATAGCATGAGTGTGAGAAAGGTTCTTGATCCATATGCCCAAGGAT 1279
|
|
|
Db 341 CymethylserinetserrapvalArgArgValPheuleuclProtyralahlsArgAsp 360
|
|
|
QY 1280 GAGCCCAATCATCAATGGTGGCTTATCAAGAAAGAGTCCCTATCCAGGCCAGAACT 1339
|
|
|
Db 361 GlyProasmtyrGlntrpValProtyrGlnGlyArgValProtyrProArgProGlyThr 380
|
|
|
QY 1340 TGTCACAGAAACATTTGGTGGTTTGAATCTCAAGAACGCTTCTGTATGATGTATA 1399
|
|
|
Db 381 CysProserlythrPheGlyGlyPheAspserThrylsAspneuProAspAspValile 400
|
|
|
QY 1400 ACCTTTCAAGAGATCATCCAGCATGTACATCCAGTGTCTTCAATCAATCCGCCA 1459
|
|
|
Db 401 ThrPhealIargerhserProIametyrAsnProValPheProhetaAsnAsnArgPro 420
|
|
|
QY 1460 ATATGTATCAAAAACGATGTAAATTTATCAATTTACACAAATTTGTCGATGCCAGATGAT 1519
|
|
|
Db 421 IleValIlelystrAspValIasnTyrglnPheThrcGlnIleValIasparGValAsp 440
|
|
|
QY 1520 GCAGAGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1579
|
|
|
Db 441 AlaGluAspGlyGlnTyraAspValIamPheIleGlyThrAspValGlyThrValIleuLys 460
|
|
|
QY 1580 GTAGTTTCAATTCCTAAGAGAGATGGTATGATTTAGAAAGAGTTCGTGTAAGAAATG 1639
|
|
|
Db 461 ValValserIleProlysglnThrIprTyraAspLeuGlnValIleuLeuGlnIumet 480
|
|
|
QY 1640 ACAGTTTTTGGGAACCGACTGCTTATTCAGCAATGAGGCTTTCACATAAGCAACA 1699
|
|
|
Db 481 ThrValPheargGlnProthrAlaIleSerAlametylIeuserThrylsGlnGlnGln 500
|
|
|
QY 1700 CTATATATGTTGTTCAACGCGTGGGTTGCCACTCCCTTACACCGGTGTATTTATTA 1759
|
|
|
Db 501 LeuTyrlleGlyserThrAlaGlyValAlaGlnLeuProleuhsArgCysAspIleTy 520
|
|
|
QY 1760 GGGAAAGCGTGTGCTAGTGTGCTGCTGCCCGCAAGCCCTTACGTGCTGGATGGTTC 1819
|
|
|
Db 521 GlylyAlaIcysAlaGlyCysGlyleuAlaArgAspProtyrGlyAlaItrAspGlySer 540
|
|
|
QY 1820 GCATGTTCTGCTATTTTCCCACTGCAAGAGACGCAACAGACGACAGATATTAAGAAAT 1879
|
|
|
Db 541 AlaCysSerArgGlyrPheProthrAlaIyAsArgThrArgArgGlnAspIleArgAsn 560
|
|
|
QY 1880 GAGAGCCCACTGACTGCTGCTGCTGCTTACACCATGATTAACAGAGCCCAACGCTT 1939
|
|
|
Db 561 GlyAspProleuthrAlaCysSerAspLeuAlaAspAsnAlshlsGlyAspSerPro 580
|
|
|
QY 1940 GAAGAGAGATCATCTATGCTGTAGAGAAATAGTACATTTTGGAAATGCAAGTCCGAG 1999
|
|
|
Db 581 GlnGluArgIleIleTyrglyValGluAsnSerSerThrPheleuGlyCysSerProLys 600
|
|
|
QY 2000 TGCAGAGAGCGCTGTCTATTGGCAATTCAGAGCGCAATGAAAGCAAGAAAGAAAG 2059
|
|
|
Db 601 SerGlnArgAlaIeuValTyrrTrpGlnPheGlnArgAsnGlnGluArgLysGlnGlu 620
|
|
|
QY 2060 ATCAGAGTGAATATATATCATCAGAGACAGATCAAGGCTTCTGTAGTATGATCAAA 2119
|
|
|
Db 621 IleArgValAspAspAlaIleIleArgTrnAspGlnGlyleuIeuIeuArgSerLeuGln 640
|
|
|
QY 2120 CAGAAAGATTCAGGCAATTAACCTCTGCCATGCGGTGGAACATGGTTCATCAAACTCT 2179
|
|
|
Db 641 GlnLysAspSercllysenTyrlleuCyshlsAlaValGlnIshlsGlyPheIleGlnThrleu 660
|
|
|
QY 2180 CTTAAGGTAACCTGGAAGTCATGACAGAGACATTTGGAAAGACTTCTTCAATAAGAT 2239
|
|
|
Db 661 LeuLysValIleThrleuGlnValIleAspThrGlnIshlsleuGlnGluIeuIeuIshLysAsp 680
|
|
|
QY 2240 GATGATGAGATGCGCTTAAAGCAAGAAATGCTCAATAGCATGACACTTAGCCGAAG 2299
|
|
|
Db 681 AspAspGlyAspDlySerlystrLysGlnIumetSerAsnSerIethrProserGlnLys 700

```

```

QY 2300 GTCTGATCAGAGACTTCATGAGCTCATCAACCAACCCCAATCTCAACAGATGAG 2359
|
|
|
Db 701 ValItrPtyrArgAspPheIleuclIleuAlaAsnIleAsnIleProAsnLeuAsnThrIethAspGlu 720
|
|
|
QY 2360 TTCTGTGAACAAGTTTGGAAAAGGAGCCGAAAACAAGTCCGCAAGAGCCAGACATACC 2419
|
|
|
Db 721 PheCysGlnGlnValItrPtyrAspArgLysGlnArgGlnArgGlnArgProGlyIshThr 740
|
|
|
QY 2420 CCAAGGAACAGTACAAATGGAAGCATTTACAGAAATTAAGAAAGTAAAGAGAGAG 2479
|
|
|
Db 741 ProGlyAsnSerAsnLysItrPtyrshlsleuGlnGluAsnLysLysGlyArgAsnArgArg 760
|
|
|
QY 2480 ACCCAAGATTTAGAGAGGCAACCCAGAGATGTC 2512
|
|
|
Db 761 ThrIshGluPheGluArgAlaProArgSerVal 771
|
|
|
RESULT 10
US-10-262-538-10
; Sequence 10, Application US/10262538
; GENERAL INFORMATION:
; APPLICANT: Alltalo et al
; TITLE OF INVENTION: NEUROBLIN/VEGF-C/VEGR-3 MATERIALS AND METHODS
; FILE REFERENCE: 28967/37564
; CURRENT FILING DATE: 2002-09-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
; LENGTH: 771
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-538-10

Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conserved: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-262-538-10 (1-771)
QY 200 ATGGCGCTGTTACTAGATGCTGTCTTTCTGGGAGTATTACTTACAGACAGACA 259
|
|
|
Db 1 MetGlyTrpLeuThrArgIleValCysleuThrItrpGlyValIleuLeuThrAlaArgAla 20
|
|
|
QY 260 AACATCAGATGGGAAGAACATGTGCCAAGCTGAAATTAATCTTACAAAGAAATGTTG 319
|
|
|
Db 21 AsnTyrglnAsnGlyLysAsnAsnValProArgleuLysleuSerTyrllyGluIuethleu 40
|
|
|
QY 320 GAATCAACATGTGATCATCTTCAATGGCTTGGCAACAGCTCACAGTATCATACCTTC 379
|
|
|
Db 41 GlnSerAsnAsnValIleThrPheAsnGlyleuAlaAsnSerSerTyrlshIethrPhe 60
|
|
|
QY 380 CTTTGGATGAGGAACGAGTAGGCTGTATGTTGGAGCAAGATCACATATTTTCATTC 439
|
|
|
Db 61 LeuIeuAspGlnGluArgSerArgleuTyrlaGlyAlaLysAspAsnIlePheSerPhe 80
|
|
|
QY 440 GACGTGTTAATATCAAGATTTTCAAAAAGATTGTGCGCAAGTATCTTAACACAGAGA 499
|
|
|
Db 81 AspIeuValIasnIleLysAspPheGlnLysIleValItrProValSerTyrlThrArgArg 100
|
|
|
QY 500 GATGAATGACAGTGGCTGGAAGAAACATCTGTAAGAAATGTGTAATTTATCAAGATA 559
|
|
|
Db 101 AspGlnCysLysItrpAlaGlyLysAspIleleuLysGlnCysAlaAsnPheIleLysVal 120
|
|
|
QY 560 CTTAAGGATTAATCAAGACTCATCTGACCGCTGTGGAAGAGGGGCTTTTCATCAAT 619
|
|
|
Db 121 LeuLysAlaItyrAsnGlnThrshlsleuTyrlaCysGlyThrGlyAlaPhehlsProIle 140
|
|
|
QY 620 TGCACCTTACATGAATTTGACATCATCTGAGAGCAATATTTTAACTGAGAGAACTCA 679
|
|
|

```


Db 141 CysThrTyrlleGlunlleGlyhshisProGluaspasnillepheylysleugluanser 160
 Qy 680 CATTGTGAAAACGGCGGTGGGAAGAGTGCATATGACCTTAAGCTGTGACAGACCTCTT 739
 Db 161 HispheGluasncllyargglyysersProTyraSProllysleuLeuthralasrleu 180
 Qy 740 TTAATAGATGAGAAATTATATCTGTGAACGTGACGTGATTTATGGGGCGAGACTTTGCT 799
 Db 181 LeuileaspGlygluleutyrsersGlyThrAlaAlaspheMetGlylthargaspheAla 200
 Qy 800 ATCTTCGGAACCTCTGGGACACCAACCAATCGAGACAGAGAGATATATTCAGCTGG 859
 Db 201 llepheargThrleucllyhshisshisProlleargThrGluGlnhshaspsersArgTyr 220
 Qy 860 CTCATGATCCAAAGTTCATTAGTGCCCACTCATCTCAGAGAGTGAACATCCTGAAGAT 919
 Db 221 LeuasnaspProlysheilleserAlahshleuilleserGluSeraspasnProGluasp 240
 Qy 920 GACAAAGTATACTTTTCTTCGTGAAAAATGCATATGATGAGAAACACTCTGGAAAAAGCT 979
 Db 241 AspIysValTyrPhepheheargGluasnAlaAlaspGlygluhisserGlylsala 260
 Qy 980 ACTCAGCTTGAATAGTGCATATGCAAGATGATGATCTTGGAGGCGACAGAGCTCTGTG 1039
 Db 261 ThrhshsalaarglleGlyGlnlleCysIysasnaspPheGlyglYhshargserleuVal 280
 Qy 1040 AATAAATGACAACTTCTCAAGCTGCTGATTTGCTGAGTGCAGTGCAGATCCAAATGGC 1099
 Db 281 AsnIysTyrThrThrPheleuysalaargleuilleserValProGlyProhansgly 300
 Qy 1100 ATGCACTCATTTTGTGATGACAGTGCAGATGATATCTTAATGAACCTTTAAGATCCTTAA 1159
 Db 301 lleaspThrhisPheheargIuleuglnaspValPheleuMetasnphelysaspProlys 320
 Qy 1160 AATCCAGTTGTATATGAGAGTGTTTACGACTTCCAGTAACATTTTCAAGGATCAGCCGTG 1219
 Db 321 AsnProvalValTyrcllyalPheThrThrSerSerAsnillepheylysGlyserAlaVal 340
 Qy 1220 TGTATGTATAGCATGATGATGAGAGAGGTGTCTTGTCTCAATGACCCACAGGAT 1279
 Db 341 CysMetTyrSerMetSeraspValahargaspValPheleuGlyProTyrAlahshargasp 360
 Qy 1280 GGAACCAATCATCATGAGGTGCTTATCAAGAAAGATCCCTATCCAGGCGCAGAGACT 1339
 Db 361 GlyProasnTyrGlnTyrValProTyrGlnGlyArgValProTyrProargProGlyThr 380
 Qy 1340 TGTCCACGAAACATTTGGTGTGTTTGAACCTTCAAGAGACCTTCTGATGATGTTATA 1399
 Db 381 CysProserLysThrPheGlyglYpheaspserThrIysaspLeuProaspaspValIle 400
 Qy 1400 ACCTTGCAAGAAGTATCAGCAATGATCAATCCAGTGTCTTCAATGAAACATGCGCCA 1459
 Db 401 ThrPheAlaSerSerhisProAlaMetTyrAsnProvalPheProMetasnshargPro 420
 Qy 1460 ATAGTATCAAAAACGAGTAAATTAATTAACAATTTGCTAGAACGAGTGCAT 1519
 Db 421 lleValillelyThraspValasnTyrGlnPheThrGlnilleValValasphargValasp 440
 Qy 1520 CGAAGAAGTGAACAGTATGATGTTATGTTATCGAAGACAGATGGTGGACCTTTTAA 1579
 Db 441 AlaGluaspGlyglintyraspValMetPheilleGlyThrAspValGlyThrValleuys 460
 Qy 1580 GTAGTTTCAATTCCTTAAGAGACTGCTATGATTTAGAGAGCTTCTGCTGAAGAAAG 1639
 Db 461 ValIysSerIleProlysglnThrTyrIysaspLeuGlnGluValleuLeuGlnGluMet 480
 Qy 1640 ACAAGTTTTCGGAACCGACTGCTATTTCAAGATGAGACTTTCACATAAGACGACACA 1699
 Db 481 ThrValPheargGluProThrAlaIleSerAlaMetGluLeuSerThrIysGlnGlnGln 500
 Qy 1700 CTATATATTTGTTCAACGCGTGGGTTGCCAGCTCCCTTTACACCGGTGTGATTTAC 1759
 Db 501 LeuTyrlleGlySerThrAlaGlyValAlaGlnleuProleuThisargCysaspIleTyr 520

Qy 1760 GGAAGAAGCGTGTGCTGAGTGTGCTCGCCGACAGACCTTACTGTGCTGGATGTTCT 1819
 Db 521 GlyIysAlaCysAlaGluCysCysleuAlahargaspProTyrCysAlaTyraspGlySer 540
 Qy 1820 GCATGTTCTCGCTATTTTCCACTGCAAGAGACGACAGACGACAGATATAAATAAT 1879
 Db 541 AlaCysSerArgTyrPheProThrAlaIysargTyrArgGlnAspIleargasn 560
 Qy 1880 GGAAGCCCACTGACCTGTCAGACTTCAACCAATGATTAACACAGCCGACGACCT 1939
 Db 561 GlyaspProleuThrhisCysSeraspLeuThshisAspAsnIshisGlyhisserPro 580
 Qy 1940 GAAGAGAGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1999
 Db 581 GlnGluArgIleIleTyrcllyalGlnIysnserserThrPheleuGluCysserProlys 600
 Qy 2000 TCGCAGAGACGCTGTCTTATTTGCAATTCAGAGGCGAATGAAGACGAAAGAG 2059
 Db 601 SerGlnAlaGlnAlaIleuValTyrTyrGlnPheGlnAlaGlnArgGlnGluArgGlyGln 620
 Qy 2060 ATCAGAGTGAATGATCATATCATCATGACAGATCAAGGCTTCTGCTACGTATGCTTACA 2119
 Db 621 lleargValaspaspshisIlelleargThraspGlnGlyleuLeuAspserleuGln 640
 Qy 2120 CAGAAGATTCAGCAATTAACCTCTGACCATGCGGTGGAACATGGGTTCATACAACTCTT 2179
 Db 641 GlnIysaspserGlyAsnTyrleuCyshisAlaValGlnhshisGlyPheIleGlnThrleu 660
 Qy 2180 CTTAAGGTATACCTTGAAGTCAATTGACACAGAGATTTGGAAGAACTTCTCATTAAGAT 2239
 Db 661 LeuIysValThrleuGlnValilleaspThrGlnhshisleuGlnGlnleuLeuThshysasp 680
 Qy 2240 GATGATGAGATGCTCTTCAAGCAACCAAGATGTCTCATATGATGACCTTAGCCAGAG 2299
 Db 681 AspaspGlyaspGlySerIysThrIysGlnMetSerAsnSerMetThrProserGlnIys 700
 Qy 2300 GTCTGTGACAGAGATTCATGAGCTCATCAACCAACCCCAATCTCAACAGATGATGAG 2359
 Db 701 ValTyrTyrIysaspPheMetGlnleuIleasnshisProasnleuasnThrMetaspGln 720
 Qy 2360 TTCTGTGAACAAGTTTGGAAAAAGGACCGAAGAACACAGCTGCGCAAGGCGCAGACATACC 2419
 Db 721 PheCysGlnGlnValTyrIysIysaspArgIysGlnArgGlnArgProGlyhshisThr 740
 Qy 2420 CCAAGGAACAGTAAACATGAGACGCTTACAGAAATTAAGAAAGTGAAGACGAGAG 2479
 Db 741 ProGlyAsnSerAsnIysTyrIysshisleuGlnGlnleuIysGlyArgAsnArgArg 760
 Qy 2480 ACCCAGAAATTTAGAGGGACCCCGAGAGTGC 2512
 Db 761 ThrhisGluPheGlnArgAlaProargserVal 771

RESULT 11
 US-60-323-784-164
 ; Sequence 164, Application US/60323784
 ; GENERAL INFORMATION:
 ; APPLICANT: Mikita, Thomas
 ; APPLICANT: Shiftman, Dov
 ; APPLICANT: Porter, Gordon, J.
 ; APPLICANT: Kaser, Matthew R.
 ; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
 ; FILE REFERENCE: PA-0050 P
 ; CURRENT APPLICATION NUMBER: US/60/323,784
 ; NUMBER OF SEQ. ID NOS: 186
 ; SOFTWARE: PERL Program
 ; SEQ. ID NO. 164
 ; LENGTH: 771
 ; TYPE: PRF
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature

OTHER INFORMATION: Incycle ID No: 1930967CD1
US-60-323-784-164

Alignment Scores:

Pred. No.:	0	Length:	771
Score:	771.00	Matches:	771
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	86.15%	Indels:	0
DB:	31	Gaps:	0

US-09-774-490-1 (1-2709) x US-60-323-784-164 (1-771)

```

QY 200 ATGGGCTGTTAACTAGATTTCTGCTTTTCTGGGAGATTAATTACAGCAAGACA 259
Db 1 MeGclYrIreuthrArGllleValCyLeuPhetrrpGlyValLeuLeuthrAlaAga 20
QY 260 AACTATCAGATGGAGAAACAATGTGCCAAGCTGAATTAATCTTACAAGAAATGTTG 319
Db 21 AsnTyGlnAnGlyLysAsnAenValProArgLeuLysSerTyrlYsGlnMetLeu 40
QY 320 GAATCCAAATGATGACATTTCAATGGCTGGCCAAAGCTCCACTTATCATCTTC 379
Db 41 GluSerAsnAnValIletrrPheAnGlyLeuAlaAnSerSerTyrlYrHsthrPhe 60
QY 380 CTTTGGATGAGGAACGAGTAGGCTGTATGTTGAGCAAGATCATATTTTCATTC 439
Db 61 LeuLeuAspGlnGluAAsGserAArgLeuTyValGlyAlaLysAspHisIlePheSerPhe 80
QY 440 GACCTGTTAAATCAAGATTTTCAAAAGATTTGTGGCCAGTATCTTACCCAGAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValIrrProValSerTyrlYrHrArgArg 100
QY 500 GATGAATGCAAGTGGCTGGAAGAAAGACATCTGAAAGATGTGTATTTTCAACAAGTA 559
Db 101 AspGlnLysStrpAlaGlyLysAspIleLeuLysGlnLysAlaAnPheIleLysVal 120
QY 560 CTTAAGGCATATATCAGACTCTTGTACGCTGTGGAACGGGGCTTTTCACTCAAT 619
Db 121 LeuLysAlaIeYrAnGlnIntrHnHisLeuTyrlAlaCySerGlyHrGlyAlaPheHisProIle 140
QY 620 TGCACTACATGAAATGATGACATCATCTGAGAGCAATATTTTAACTGGAGAACTCA 679
Db 141 CysThrTyrlIleGlnIleGlyHisIleAspProGlnAspAsnIlePheLysLeuGlnAsnSer 160
QY 680 CATTTGAAACCGCGCTGGGAGAGTCCATATGACCTTAAGTCTGAGAGATCCCTT 739
Db 161 HisPheGlnAsnGlyArgGlyLysSerProTyraSerProLysLeuLeuthrAlaSerLeu 180
QY 740 TTATAGATGAGAAATTAATTACTCTGAACTGCAAGTGAATTTTATGGGGCAGACTTGTCT 799
Db 181 LeuIleAspGlyGlnLeuLeuTyrlSerGlyThrIleAlaAspPheMetGlyArgAspPheAla 200
QY 800 ATTTCCGAACTTTGGGACACCAACCCCAATCAGACAGACAGACATGATTCGAGGTG 859
Db 201 IlePheArgThrLeuGlnHisIleHisIleProIleArgThrGlnGlnHisIleAspSerArgTyr 220
QY 860 CTGAATGATCCAAAGTTCATTAAGTCCCACTCATCTCAGAGAGTGCATCTCTGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisIleLeuIleSerGlnSerAspAsnProGlnAsp 240
QY 920 GACAAAGTACTTTTCTTCCTGGAATGATGAGATGAGAGACATCTGGAAGAAAGCT 979
Db 241 AspLysValIeYrThrPhePheArgGlnAsnAlaIleAspGlyGlnHisSerGlyLysAla 260
QY 980 ACTCAGCTAGAAATAGGTCAGATATGCAAGATGACTTTGGAGGGCAGAGAGCTTGCTG 1039
Db 261 ThrHisAlaArgIleGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATAATGACAACTTCTCCAAAGCTCTGATTTGCTCAGTGCAGAGTCCAAATGGC 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300

```

```

QY 1100 ATGCACTCATTTTGTAGTAAGTCAAGATGATTAATCTTAATGAATTTAAAGCTCTAA 1159
Db 301 IleAspThrHisPheAspGlnLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
QY 1160 AATCCAGTTGTATATGAGTGTTTTACGACTTCCAGTAACATTTTCAAGGATCAGCCGTG 1219
Db 321 AsnProValValTyrlYalPheThrTrpSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATGTATAGATGATGATGATGAGAGAGGTTCCTTGTCATATGCCACAGGGAT 1279
Db 341 CysMetTyraSerMetSerAspValAlaArgValPheLeuGlyProTyrlAlaHisArgAsp 360
QY 1280 GACCCCAATCATATGAGTGGCTTATCAAGAGAGTCCCTATCCAGGCGCAGAGACT 1339
Db 361 GlyProAsnTyrlGlnTrpValProTyrlGlnIleArgValProTyraProArgProGlyThr 380
QY 1340 TGTCCAGCAAAACATTTGGTGGTTTGACTCTCAAGAGACCTTCTGATGATGTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerTrpLysAspLeuProAspAspValIle 400
QY 1400 ACCTTTGCAAGAGTATCAGCAATGTACATTCAGAGTTCCTATGAACCAATGCCCA 1459
Db 401 ThrPheAlaArgSerHisPheAlaMetTyraAspProValPheProMetAsnAsnPro 420
QY 1460 ATAGTATCAAAACGAGATGTAATTAATTAATTAACAATTTGCTAGACCGAGTGAAT 1519
Db 421 IleValIleLysThrAspValAsnTyrlGlnPheTrnGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAAGATGACAGTATGATGTATGTTATCCGAAACAGATGTTGGACCGTCTTAA 1579
Db 441 AlaGlnAspGlyGlnTyraPheValMetPheIleGlyThrAspValGlyThrValLeuLys 460
QY 1580 GTAGTTTCAATTCCTAAGAGACTGTGTATGATTTAGAAGAGCTTCGTGGAGAAATG 1639
Db 461 ValValSerIleProLysGlnTrpTyraAspLeuGlnIleValLeuLeuGlnLeuMet 480
QY 1640 ACAGTTTTTCGGAAACCGACTGTATTTCAAGATGAGCTTTCACATAGCAGCAACA 1699
Db 481 ThrValPheArgGlnProThrAlaIleSerAlaMetGlnLeuSerThrLysGlnGln 500
QY 1700 CTATATATTGTTCAACGAGCTGGGTTGGCCAGCTTCCTTTACACCGGTGTATTTAC 1759
Db 501 LeuTyrlIleGlySerThrAlaGlyAlaAlaGlnLeuProLeuHisArgCysAspIleTy 520
QY 1760 GGAAGAGGCTGTGAGTGTGCTCGCCGCGCAGACCTTATGTGCTTGAGAGTGTCT 1819
Db 521 GlyLysAlaCysAlaGlnLysCysLeuAlaArgAspProTyraAlaIrrAspGlySer 540
QY 1820 GCATGTTCTCGTATTTTCCACTGCAAGAGACGCAACAAGACCAAGATTAAGAAAT 1879
Db 541 AlaCysSerArgTyraPheProThrAlaLysArgAlaGlnThrArgGlnAspIleArgAsn 560
QY 1880 GAGAGCCCACTGACTCATCTGTTCAACTTACCAATGATATACCATGCGCACAGCCTT 1939
Db 561 GlyAspProLeuThrHisIeCysSerAspLeuHisIleAspAsnHisIleGlyHisSerPro 580
QY 1940 GAAAGAGAAATCATGTATGCTGTAGAGATAGTACAGCAATTTTGAATGAGCTCGAAG 1999
Db 581 GlnGlnLysGlyIleIeTyrlGlyAlaGlnAsnSerSerTrpPheLeuGlnCysSerProLys 600
QY 2000 TCGCAGAGACGCTGGTCTATTGGCAATTCAGAGCGCAATGAAGGGAAGAAAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrlTrpGlnPheGlnAlaGlnAsnGlnGlnArgLysGln 620
QY 2060 ATCAGAGTGGATGATCATATCATTCAGACAGATCAAGGCTTCTGCTACGTAGTCAAA 2119
Db 621 IleArgValAspAspHisIleIleArgThrArgGlnGlyLeuLeuLeuAspSerLeuGln 640
QY 2120 CAGAGGATTCAGGGAATTAATCTCTGCAATGGGAGGAGCAAGGGTTCATACAACTCTT 2179
Db 641 GlnLysAspSerGlyAsnTyrlLeuTyrlCysHisAlaValGlnHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAACTCTGAAGTCAATTGACACAGAGCAATTTGGAAGAACTTCTTCATAAAGAT 2239

```

```

Db      661  LeuLVsValThleuLnuValIleAspThrGlnHsleuGlnleuLeuHslyAsp 680
QY      2240  GATGATGGAGATGGCTCTAAGACCAAGAAATGTCCAAATGACATGACACTGACCGCAAG 2299
Db      681  AspaPGLysAspGlySerIsthrIysGlnleuSerAsnSerMetThrProSerGlnLys 700
QY      2300  GTCTGTACAGAGACTTCATGAGACTCATCAACCAACCCCAATCTCAACAGATGATGAG 2359
Db      701  ValTPlyrArgAspPheMetGlnleuIleAsnHsIProAsnleuAsnThrMetAspGlu 720
QY      2360  TTCTGTGAACAAGTTTGGAAAAGGACCCAAAACAAGTCGGCAAAAGCCAGACATACC 2419
Db      721  PheCySGlnGlnValTPlysArgAspArgIysGlnArgGlnArgProGlnIysIsthr 740
QY      2420  CGAGGAACAGATGACAAATGGAAGCACTTACAAAGAAATTAAGAGCTGAAGAACGAGAG 2479
Db      741  ProGlnAsnSerAsnIsthrIysIsthrIysleuGlnleuLysleuGlyArgAsnArg 760
QY      2480  ACCCAGATTTGAGAGGCGACCCAGAGTGTCT 2512
Db      761  ThrHsGlnPheGlnArgIleProArgSerVal 771

```

RESULT 12

```

US-09-791-537-32993
; Sequence 32993, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Dancet, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791, 537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 32993
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-32993

```

Alignment Scores:

```

Pred. No.: 2,74e-280 Length: 287
Score: 287.00 Matches: 287
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 32.07% Indels: 0
DB: 22 Gaps: 0

```

US-09-774-490-1 (1-2709) x US-09-791-537-32993 (1-287)

```

QY      1652  GAACGACTGCTATTTCAGCAATGAGCTTTCACCTAAGCAGCAACAATATATATGCT 1711
Db      1  GluProThrAlaIleSerIleMetGlnleuSerThrIysGlnGlnleuIsthrIleGly 20
QY      1712  TCAAGCGCTGGGCTGCCAGCTCCCTTTACACCGGTGTGATTTTACGGGAAAGCGTGT 1771
Db      21  SerThrAlaGlyValAlaGlnleuProleuHsIsthrCyAspDlleIsthrIysAlaCys 40
QY      1772  GGTGAGTGTGGCTGGCCGAGACCTTACTGTGCTGGAGTGGTGTCTGCAATGTTCTGCC 1831
Db      41  AlaGlnCySGlyLeuAlaIsthrAspProIsthrCyAspIsthrAspGlySerIsthrIys 60
QY      1832  TATTTTCCCACTGCAAGAGAGCGACACAGACGACAAATATTAAGAAATGAGACCCACTG 1891
Db      61  TyrPheProThrAlaIsthrArgIsthrIsthrArgGlnAspIleArgAsnIsthrAspProleu 80
QY      1892  ACTCATCTGTTCAGACTTACACATGATATATCAACATGGCCAGACCCCTGAAGAAGAAC 1951
Db      81  ThrHsCysSerAspPheuHsIsthrAspAsnHsIsthrIsthrProGlnGlnIsthrIsthr 100

```

```

QY      1952  ATCTATGCTGTAGAGAAATAGTACACATTTTGAATGACAGTCCGAGAGTCCGAGAGCG 2011
Db      101  IleTyrGlnValGlnAsnSerSerThrPheleuGlnIsthrProIsthrSerGlnIsthrGln 120
QY      2012  CTGGTCTATTTGGCAATTTCCAGAGCGGAATGAGAGCCGAAAAGAGAGATCGAGTGGAT 2071
Db      121  LeuValIsthrTPGlnPheGlnIsthrArgAsnGlnIsthrIsthrIsthrIsthrIsthr 140
QY      2072  GATCATATCATGAGAGACAGATCAAGACCTTCTGCTACAGTCTCAACAGAGAGATTCA 2131
Db      141  AspHsIsthrIsthrIsthrAspGlnIsthrIsthrIsthrIsthrIsthrIsthrIsthr 160
QY      2132  GGCATTAACCTCTGCGCATCGCGTGGACATGGGTTTCATACAACTTTCTTAAGGTAAAC 2191
Db      161  GlnAsnIsthrIsthrIsthrIsthrIsthrIsthrIsthrIsthrIsthrIsthrIsthr 180
QY      2192  CTGGAAGTCTTGTACACAGAGATTTTGAAGAATCTTCTCATTAAGATGATGATGAT 2251
Db      181  LeuGlnValIleAspThrGlnHsleuGlnleuLeuHsIsthrAspAspAspGlyAsp 200
QY      2252  GGCTCTAAGACCAAGAAATGTCCATAGACATGACACCTAGCAGAGAGTCTGTGTAACA 2311
Db      201  GlySerIsthrIsthrIsthrIsthrIsthrIsthrIsthrIsthrIsthrIsthrIsthr 220
QY      2312  GACTTCATGCAAGCTCATGACACCAACCAATCTCAACAGATGATGATGATGATGATGAT 2371
Db      221  AspPheMetGlnleuIleAsnHsIsthrAsnleuAsnIsthrIsthrIsthrIsthrIsthr 240
QY      2372  GTTTGGAAAAGGAGACCGAAAACAAGCTGGGCAAGGCGCAAGATATCCCGAGGAACGT 2431
Db      241  ValTPlysArgAspArgIsthrIsthrIsthrIsthrIsthrIsthrIsthrIsthrIsthr 260
QY      2432  AACAAATGAGACACTTCAAGAAATTAAGAGTGAAGAACAGAGAGCCAGCAATTT 2491
Db      261  AsnIsthrIsthrIsthrIsthrIsthrIsthrIsthrIsthrIsthrIsthrIsthrIsthr 280
QY      2492  GAGAGGAGACCCAGAGAGTGTCT 2512
Db      281  GluArgIleProArgSerVal 287

```

RESULT 13

```

US-09-724-676-66689
; Sequence 66689, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724, 676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 66689
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-724-676-66689

```

Alignment Scores:

```

Pred. No.: 4,45e-263 Length: 272
Score: 270.00 Matches: 270
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 30.17% Indels: 0
DB: 21 Gaps: 0

```

US-09-774-490-1 (1-2709) x US-09-724-676-66689 (1-272)

```

QY      200  ATGGGCTGGTTAACTAGAGATGTCTGCTTTCTGGGAGTATTAATTACAGACAGAGCA 259
Db      1  MetGlyTPleuThrIsthrIsthrIsthrIsthrIsthrIsthrIsthrIsthrIsthrIsthr 20
QY      260  AACTATCAGAAATGGAGAACAAATGTGCCAAGGCTGAATTAATCTTACAAAGAAATGTTG 319

```

```

Db      21 AsnlyrGlnaAnglylysaAsnValProArgleuylseuSerTyrlYsglUmeLeu 40
QY      320 GAATCCAACAATGATGATCACTTTCATATGGCTGGCCCAACAGCTCCAGTTATCATCTTC 379
Db      41 GluSerAsnAsnValIlethrPheAnglyLeuAlaAsnSerSerSerTyrlYsrhPhe 60
QY      380 CTTTGGATGAGAACGAGTAGGCTGTATGTTGAGCAAGATCATATTTTTCATTC 439
Db      61 LeuLeuAspGluGluArgSerArgLeuTyrlValGlyAlaYsaPheSrllePheSerPhe 80
QY      440 GACCTGTTAATATCAAGATTTTCAAAAGATTTGTGGCCAGTATTTACACCAAGA 499
Db      81 AspleuValAsnIleuYsaPheGlnlyValIleValIleProValSerTyrlYsrArg 100
QY      500 GATGAATGCAAGTGGGCTGGAAGAGACATCTGAAAGATGTGTAATTTTCAACAGTA 559
Db      101 AspGluCysLeuStrpAlaGlyYsaPrlleuLeuYsaGlyCysAlaAsnPheIleYsaVal 120
QY      560 CTTAAGCATATATATCAAGACTCATCTGTACGCTGTGGAACGGGGCTTTTTCATCCAA 619
Db      121 LeuYsaAlaTyrlaAnglnIthrIleuTyrlaCysGlyThrGlyAlaPheIleSproIle 140
QY      620 TGCACTCACTATGAAATTTGACATCATCTGAGACATATTTTAACTGGAGAACTCA 679
Db      141 CysThrTyrlleGlnIleGlyIleGlyIleSproGlnAspAsnIlePheYsaLeuGlnAsnSer 160
QY      680 CATTTGAAAACGGCGCTGGGAAGAGTCATATGACCTTAAGCTGAGCAGCATCCCTT 739
Db      161 HisPheGlnAsnIleYsaGlyYsaSerProTyraSprolyseuLeuIthrAlaSerLeu 180
QY      740 TTAATAGATGAGAAATTTATCTGTGAACCTGAGCTGATTTTATGGGCGAGACTTGGCT 799
Db      181 LeuIleAspGlyGlnLeuTyrlSerGlyThrAlaIleAspPheMetGlyArgAspPheAla 200
QY      800 ATCTTCGGAACCTTTGGGCAACCAACCATAGACAGACAGACATGATTTCCAGTGG 859
Db      201 IlePheArgIleuGlnIleGlyIleSproIleArgIleGlnIleSproIleArgIle 220
QY      860 CTCATGATCCAAAGTTGATTAAGTCCACCTCATCTGACAGAGAGTGAACATCCTGAAGAT 919
Db      221 LeuAsnAspProlysePheIleSerIleAlaIleSeriIleSerGlnSerAspAsnProGlnAsp 240
QY      920 GACAAAGTATCTTTTCTTCGCTGAATAATGCAATAGATGAGAACATCTGGAAGAGCT 979
Db      241 AspIlyValTyrlPhePhePheArgGlnAsnAlaIleAspGlyGlnHisSerGlyYsaAla 260
QY      980 ACTCAGCTAGAAATAGTACAGATATGCAAG 1009
Db      261 ThrHisAlaArgIleGlyGlnIleCysIys 270

RESULT 14
US-09-724-676A-66689
; Sequence 66689, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: Compugen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 Compugen
; CURRENT APPLICATION NUMBER: US/09/724.676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 66689
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-724-676A-66689

Alignment Scores:
Pred. No.: 4,45e-263 Length: 272
Score: 270.00 Matches: 270
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 30.17% Indels: 0

```

```

DB:      21 Gaps: 0
US-09-774-490-1 (1-2709) x US-09-724-676A-66689 (1-272)
QY      200 ATGGGCTGGTTAACTAGAGATGTCTGTCTTTTCTGGGAGATTAATTAACAGCAAGACA 259
Db      1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValIleuLeuThrAlaArgAla 20
QY      260 ACTCTACAAAGGAGAAACAATGTCACAGGCTGAAGATTTCTCTACAAAGAAATGTTG 319
Db      21 AsnlyrGlnaAnglylysaAsnValProArgleuylseuSerTyrlYsglUmeLeu 40
QY      320 GAATCCAACAATGATGATCACTTTCATATGGCTGGCCCAACAGCTCCAGTTATCATCTTC 379
Db      41 GluSerAsnAsnValIlethrPheAnglyLeuAlaAsnSerSerSerTyrlYsrhPhe 60
QY      380 CTTTGGATGAGAACGAGTAGGCTGTATGTTGAGCAAGATCATATTTTTCATTC 439
Db      61 LeuLeuAspGluGluArgSerArgLeuTyrlValGlyAlaYsaPheSrllePheSerPhe 80
QY      440 GACCTGTTAATATCAAGATTTTCAAAAGATTTGTGGCCAGTATTTACACCAAGA 499
Db      81 AspleuValAsnIleuYsaPheGlnlyValIleValIleProValSerTyrlYsrArg 100
QY      500 GATGAATGCAAGTGGGCTGGAAGAGACATCTGAAAGATGTGTAATTTTCAACAGTA 559
Db      101 AspGluCysLeuStrpAlaGlyYsaPrlleuLeuYsaGlyCysAlaAsnPheIleYsaVal 120
QY      560 CTTAAGCATATATATCAAGACTCATCTGTACGCTGTGGAACGGGGCTTTTTCATCCAA 619
Db      121 LeuYsaAlaTyrlaAnglnIthrIleuTyrlaCysGlyThrGlyAlaPheIleSproIle 140
QY      620 TGCACTCACTATGAAATTTGACATCATCTGAGACATATTTTAACTGGAGAACTCA 679
Db      141 CysThrTyrlleGlnIleGlyIleGlyIleSproGlnAspAsnIlePheYsaLeuGlnAsnSer 160
QY      680 CATTTGAAAACGGCGCTGGGAAGAGTCATATGACCTTAAGCTGAGCAGCATCCCTT 739
Db      161 HisPheGlnAsnIleYsaGlyYsaSerProTyraSprolyseuLeuIthrAlaSerLeu 180
QY      740 TTAATAGATGAGAAATTTATCTGTGAACCTGAGCTGATTTTATGGGCGAGACTTGGCT 799
Db      181 LeuIleAspGlyGlnLeuTyrlSerGlyThrAlaIleAspPheMetGlyArgAspPheAla 200
QY      800 ATCTTCGGAACCTTTGGGCAACCAACCATAGACAGACAGACATGATTTCCAGTGG 859
Db      201 IlePheArgIleuGlnIleGlyIleSproIleArgIleGlnIleSproIleArgIle 220
QY      860 CTCATGATCCAAAGTTGATTAAGTCCACCTCATCTGACAGAGAGTGAACATCCTGAAGAT 919
Db      221 LeuAsnAspProlysePheIleSerIleAlaIleSeriIleSerGlnSerAspAsnProGlnAsp 240
QY      920 GACAAAGTATCTTTTCTTCGCTGAATAATGCAATAGATGAGAACATCTGGAAGAGCT 979
Db      241 AspIlyValTyrlPhePhePheArgGlnAsnAlaIleAspGlyGlnHisSerGlyYsaAla 260
QY      980 ACTCAGCTAGAAATAGTACAGATATGCAAG 1009
Db      261 ThrHisAlaArgIleGlyGlnIleCysIys 270

RESULT 15
US-09-864-761-47112
; Sequence 47112, Application US/09864761
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wenpeng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecmca-X-1
; CURRENT APPLICATION NUMBER: US/09/864.761
; CURRENT FILING DATE: 2001-05-23

```

;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 47112
;; LENGTH: 151
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC006322.2
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.73
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.55
;; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALU2 2.00e-88
US-09-864-761-47112

Alignment Scores:
Pred. No.: 1,35e-142 Length: 151
Score: 151.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.87% Indels: 0
DB: 23 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-864-761-47112 (1-151)

QY 2060 ATCAGAGTGCATGATCATATCATGACAGACAGATCAAGCCCTTCTGCTAGTACTACAA 2119
DB 1 IleaTgValaSpAphIslelleIeaTgThraSpGInglYleuLeuLeuAArgSerleuGIn 20
QY 2120 CAGAGGATTCGAGCAATTAATCTGCGCCATGGCGGAGGAACAGGGTTCATTAACACTCTT 2179
DB 21 GInlySaSpSerGlyAsnlyrleuCyshIsaIaValaGluhSeglyPheIleGInThreU 40
QY 2180 CTTAAGGTAACTCGAAGTCAATGACACAGACAGCAATTTGGAAGAATCTTCTATTAAGT 2239
DB 41 leuYsValThrlleuGluValIleAspThrGluhIsleuGInleuLeuAArgSerleuGIn 60
QY 2240 GATGATGAGATGGCTCTAAGACCAAGAAATGTCATATGATGATGACCTTACCCAGAG 2299
|||||

DB 61 AspaSpGlyAspGlySerlySerlyThrlySglUwetSerAenSerMetThrProSerGInlyS 80
QY 2300 GCTGGTACAGAGCATTCATGACGCTCATCAACCAACCAATCTCAACAGATGATGAG 2359
DB 81 ValTrpYrArgAspPheMetCInleuIleAsnHsProAsnLeuAAsnThrMetAspGlu 100
QY 2360 TTCTGTGAACAATTGGAGAAAGGAGCAACCAACGTCGCAAAAGGCACAGACATACC 2419
DB 101 PheCySglUglInValTrpLySaArgSpArTgysGInAArgGlnAArgProGlyhSthr 120
QY 2420 CCAGGAACAGTAAACAATGCAAGCATTAACAAGAAATAGAAAGGTAGAAACAGAGG 2479
DB 121 ProGlyAenSerAsnlySerlyPlyshIsleuGInGluAenlySlySglYArgSnaArgYr 140
QY 2480 ACCGAGATTTGAGAGGAGCCGAGAGTGC 2512
DB 141 ThrhSglUfpheGluArgAlaProArgSerVal 151

RESULT 16
US-10-203-135-32595
;; Sequence 32595, Application US/10203135
;; GENERAL INFORMATION:
;; APPLICANT: Molecular Dynamics, Inc.
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wenheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: PB 0004 WO 5
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US/10/203,135
;; PRIOR FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; NUMBER OF SEQ ID NOS: 37012
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 32595
;; LENGTH: 151
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC006322.2
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.73
;; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALU2 2.00e-88
US-10-203-135-32595

Alignment Scores:
Pred. No.: 1,35e-142 Length: 151
Score: 151.00 Matches: 151
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 16.87% Indels: 0
DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-203-135-32595 (1-151)

QY 2060 ATCAGAGTGCATGATCATATCATGACAGACAGATCAAGCCCTTCTGCTAGTACTACAA 2119
DB 1 IleaTgValaSpAphIslelleIeaTgThraSpGInglYleuLeuLeuAArgSerleuGIn 20
|||||


```

QY 1229 ACCATGATGATGTGAGAGGGTGTCTTGTGCTATATGCCACAGGATGAGCCCAAC 1288
    |||
    344 SerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgSpGlyProAsn 363
QY 1289 TATCATGGGTCCTTATATCAGAGAGTCCCTATCCAGCCGCAAGAACTGTCCACG 1348
    |||
    364 TyrGlnTyrPValProTyrGlnGlyArgValProTyrProArgProGlyThrCysProSer 383
QY 1349 AAAACATTTGGTGTGCTTACTCTCAAGAGACCTTCGATGATGATTAACCTTTGCA 1408
    |||
    384 LysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIleThrPheAla 403
Db 1409 AGAAGTCATCCAGCCATGATCAATCCAGTGTTCCT 1444
    |||
    404 ArgSerHisProAlaMetCysTrpAsnProValPhePro 415

RESULT 19
US-09-791-537-101922
; Sequence 101922, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Dabe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 101922
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-101922

Alignment Scores:
Pred. No.: 6, 47e-83 Length: 772
Score: 92.00 Matches: 92
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.28% Indels: 0
Gaps: 0
DB: 22

US-09-774-490-1 (1-2709) x US-09-791-537-101922 (1-772)

QY 1169 GTATATGAGTGTATTACGACTTCCAGTAACTTTTCAGAGATCAGCCGTGTATGAT 1228
    |||
    324 ValTyrGlyValPheThrThrSerSerAsnIlePheLeuGlySerAlaValCysMetCys 343
QY 1229 AGCATGATGATGTGAGAGGGTGTCTTGTGCTATATGCCACAGGATGAGCCCAAC 1288
    |||
    344 SerMetSerAspValArgValPheLeuGlyProTyrAlaHisArgSpGlyProAsn 363
QY 1289 TATCATGGGTCCTTATATCAGAGAGTCCCTATCCAGCCGCAAGAACTGTCCACG 1348
    |||
    364 TyrGlnTyrPValProTyrGlnGlyArgValProTyrProArgProGlyThrCysProSer 383
Db 1349 AAAACATTTGGTGTGCTTACTCTCAAGAGACCTTCGATGATGATTAACCTTTGCA 1408
    |||
    384 LysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIleThrPheAla 403
QY 1409 AGAAGTCATCCAGCCATGATCAATCCAGTGTTCCT 1444
    |||
    404 ArgSerHisProAlaMetCysTrpAsnProValPhePro 415

RESULT 20
US-09-791-537-62757
; Sequence 62757, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Dabe, Derek

```

```

; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 62757
; LENGTH: 666
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-62757

Alignment Scores:
Pred. No.: 8, 67e-73 Length: 666
Score: 82.00 Matches: 82
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9.16% Indels: 0
Gaps: 0
DB: 22

US-09-774-490-1 (1-2709) x US-09-791-537-62757 (1-666)

QY 899 GAGAGTGAACAATCTGAAGATGACAAAGTATATCTTTCTCCGTGAAGAAATGCAATAGAT 958
    |||
    128 GluSerAspAsnProGluAspAspLysValTyrPhePhePheArgGluAsnAlaIleAsp 147
QY 959 GGAGAACACTCTGGAAAGCTACTACGCTAGATAGATGATGATGATGATGATGATGAT 1018
    |||
    148 GlyGluHisSerGlyLysAlaThrHisAlaArgIleGlyGlnIleCysLysAsnAspPhe 167
QY 1019 GGAGGCGACAGAGCTGCTGAATTAATGACAAATCTCCAAAGCTGCTGATTTGC 1078
    |||
    168 GlyGlyHisArgSerLeuValAsnLysTyrThrThrPheLeuLysAlaArgLeuIleCys 187
Db 1079 TCAGTCCAGGTCCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1138
    |||
    168 SerValProGlyProAsnGlyIleAspThrHisPheAspGluLeuGlnAspValPheLeu 207
QY 1139 ATGAAAC 1144
    |||
    208 MetAsn 209
Db 1139 ATGAAAC 1144
    |||
    208 MetAsn 209

RESULT 21
US-09-262-167-32
; Sequence 32, Application US/09262167A
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; APPLICANT: Lofton-Day, Catherine B.
; TITLE OF INVENTION: HUMAN SEMAPHORIN ZSMF-7
; FILE REFERENCE: 97-59
; CURRENT APPLICATION NUMBER: US/09/262,167A
; EARLIER FILING DATE: 1999-03-03
; EARLIER APPLICATION NUMBER: 60/076,611
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-262-167-32

Alignment Scores:
Pred. No.: 9, 35e-70 Length: 772
Score: 79.00 Matches: 79
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8.83% Indels: 0
Gaps: 0
DB: 16

US-09-774-490-1 (1-2709) x US-09-262-167-32 (1-772)

```


Db 41 IIEVALILEYsthrspValenTyrgInPheThGlnIleValaspArgValasp 60
QY 1520 GCAGAGATGACGACGATGATGTTATGTTATCGGACGA 1558
Db 61 AIAGIAspGIyGlnTyAspValMetPheIleGIyThr 73
RESULT 25
US-09-791-537-111919
; Sequence 111919, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danczer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patent version 3.0
; SEQ ID NO 111919
; LENGTH: 772
; TYPE: PRT
; ORGANISM: Gallus gallus
US-09-791-537-111919
Alignment Scores:
Pred. No.: 1.69e-48 Length: 772
Score: 58.00 Matches: 58
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.48% Indels: 0
DB: 22 Gaps: 0
US-09-774-490-1 (1-2709) x US-09-791-537-111919 (1-772)
QY 971 GGAAGAGCTACTCAGCTAGATAGTACATATGACAGATGACCTTTGGAGGACAGA 1030
Db 258 GilyysaIaThrhIsaIaArgIleGIyGlnIleCyblyAsrhAspPheGIyGlnIsArg 277
QY 1031 AGCTGGTGATTAATGACAAACATTCTCTCAAGCTCGTGTGATTTGCTCAGTGCAGGT 1090
Db 278 SerLeuValaenlystrPthrPheulysaIaArgLeuIleCysserValProGIy 297
QY 1091 CCAATGGATGACATCTATTGATGATGACGACAGATGATTTCTTAATGAC 1144
Db 298 ProaenglyIleAspThrhIsaPheAspGluLeuGlnAspValPheLeuMetAsn 315
RESULT 26
US-60-181-996-1113
; Sequence 1113, Application US/60181996
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CLO000238
; CURRENT APPLICATION NUMBER: US/60/181,996
; CURRENT FILING DATE: 2000-02-11
; NUMBER OF SEQ ID NOS: 1888
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1113
; LENGTH: 76
; TYPE: PRT
; ORGANISM: HUMAN
US-60-181-996-1113
Alignment Scores:
Pred. No.: 2.03e-47 Length: 76
Score: 57.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 6.37% Indels: 0
DB: 31 Gaps: 0
US-09-774-490-1 (1-2709) x US-60-181-996-1113 (1-76)
QY 1388 GATGATGTTATACCTTTGCAAGAGATCATCCAGCATGACATTCAGTGTTCCTANG 1447
Db 17 AspaPValIleIthrPheAlaArgSerhIsProIaMetTyAsnProValPheProMet 36
QY 1448 AACATGCCCAATGTGATGATCAAAACGATGTAATTTATCAATTATCAAAATGTGCGTA 1507
Db 37 AsnaAsnArgProIleValIleIlystrAspValaenTyrgInPheThrhGlnIleVal 56
QY 1508 GACCGAGTGATGACAGAGATGACGATGATGTTATGTTATCGGACGA 1558
Db 57 AspaArgValaAspIaGlnAspGIyGlnTyAspValMetPheIleGIyThr 73
RESULT 27
US-10-029-386-30748
; Sequence 30748, Application US/10029386
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; FILE REFERENCE: A60MICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30748
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR7.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.59
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 8.00e-29
US-10-029-386-30748
Alignment Scores:
Pred. No.: 2.25e-44 Length: 54
Score: 54.00 Matches: 54
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.03% Indels: 0
DB: 26 Gaps: 0
US-09-774-490-1 (1-2709) x US-10-029-386-30748 (1-54)
QY 1691 CAGCAACATATATATGTTCAACGCGTGGGTTGCCAGCTCCCTTTACACCGGTGT 1750
Db 1 GlnGlnGlnLeuTyrlleGIySerThraIaGIyValaIaGlnLeuProLeuHIsArgCy 20
QY 1751 GATATTTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGACAGACCTTACTGTGCTTGG 1810
Db 21 AspiIeTyrgIlyValaCyalaGluCyCysLeuAlaArgAspProTyCyalaItrp 40
QY 1811 GATGTTTGCATGTTCTGCTATTTTCCCACTGCAAGAGA 1852
Db 41 AspGIySerAlaCySerArgTytrPheProThraIaIysArg 54
RESULT 28
PCT-US01-00663-33541
; Sequence 33541, Application PC/TUS0100663
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.

```
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 7
CURRENT APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38837
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 33541
LENGTH: 57
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALU 5.00e-25
PCT-US01-00663-33541
```

Alignment Scores:

Pred. No.:	2,316-43	Length:	57
Score:	53.00	Matches:	53
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.92%	Indels:	0
DB:	1	Gaps:	0

US-09-774-490-1 (1-2709) x PCT-US01-00663-33541 (1-57)

```
QY 311 GAAATGTTGAATCCAAACATGATCATCTTCAATGCTTGCGCAACAGCTCCAGTTAT 370
DB 5 GUMETLEUGLUSERSHNSNVAILLETHRPHEANGLYLEULALASNSERSERTYR 24
QY 371 CATACCTCTTTGGATGAGGAACGAGTAGGCTGTATGTTGGACCAAGATCACA 430
DB 25 HSTHPhelenuhsplugluArGserArgLeuTYrValGlyAlaYshpHisile 44
QY 431 TTTTCATTCGACCTGTTAATATCAAGATTTTCAAAAG 469
DB 45 PheSerPheAspLeuValAsnIleYshpPheGlnIlyS 57
```

RESULT 29

```
US-09-864-761-44075
Sequence 44075, Application US/09864761
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecm1ca-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
```

```
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
SEQ ID NO 44075
LENGTH: 57
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALU 5.00e-25
US-09-864-761-44075
```

Alignment Scores:

Pred. No.:	2,316-43	Length:	57
Score:	53.00	Matches:	53
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.92%	Indels:	0
DB:	23	Gaps:	0

US-09-774-490-1 (1-2709) x US-09-864-761-44075 (1-57)

```
QY 311 GAAATGTTGAATCCAAACATGATCATCTTCAATGCTTGCGCAACAGCTCCAGTTAT 370
DB 5 GUMETLEUGLUSERSHNSNVAILLETHRPHEANGLYLEULALASNSERSERTYR 24
QY 371 CATACCTCTTTGGATGAGGAACGAGTAGGCTGTATGTTGGACCAAGATCACA 430
DB 25 HSTHPhelenuhsplugluArGserArgLeuTYrValGlyAlaYshpHisile 44
QY 431 TTTTCATTCGACCTGTTAATATCAAGATTTTCAAAAG 469
DB 45 PheSerPheAspLeuValAsnIleYshpPheGlnIlyS 57
```

RESULT 30

```
US-10-182-993-32519
Sequence 32519, Application US/10182993
```

```
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 2
CURRENT APPLICATION NUMBER: US/10/182,993
CURRENT FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 37811
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 32519
LENGTH: 57
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
FEATURE:
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
FEATURE:
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUATION = 5.00e-25
US-10-182-993-32519

Alignment Scores:
Pred. No.: 2,31e-43 Length: 57
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5,92% Indels: 0
DB: 27 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-182-993-32519 (1-57)
QY 311 GAAATGTTGGAATCCAAACATGATGATCACTTCAATGCTGGCCAAAGCTCCAGTTAT 370
DB 5 GtUeUleUgUSeRshsnValllethrhphasnghlyleuAlaasnserseTyr 24
QY 371 CATACCTCTCTTTGGATGAGGAGACGAGTAGGCTGTATGTTGAGCAAGATCACATA 430
DB 25 HsthrPhelUeUaspGlUarGserArGleUyValGlyAlaYasphstle 44
QY 431 TTTTCATTGCACTGTTATATCAAGATTTTCAAAAG 469
DB 45 PheserPhespleUeUalasnleUysasphglnly 57

RESULT 31
US-10-203-134-33356
Sequence 33356, Application US/10203134
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 6
```

```
CURRENT APPLICATION NUMBER: US/10/203,134
CURRENT FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38628
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 33356
LENGTH: 57
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
FEATURE:
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
FEATURE:
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUATION = 5.00e-25
US-10-203-134-33356

Alignment Scores:
Pred. No.: 2,31e-43 Length: 57
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5,92% Indels: 0
DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-203-134-33356 (1-57)
QY 311 GAAATGTTGGAATCCAAACATGATGATCACTTCAATGCTGGCCAAAGCTCCAGTTAT 370
DB 5 GtUeUleUgUSeRshsnValllethrhphasnghlyleuAlaasnserseTyr 24
QY 371 CATACCTCTCTTTGGATGAGGAGACGAGTAGGCTGTATGTTGAGCAAGATCACATA 430
DB 25 HsthrPhelUeUaspGlUarGserArGleUyValGlyAlaYasphstle 44
QY 431 TTTTCATTGCACTGTTATATCAAGATTTTCAAAAG 469
DB 45 PheserPhespleUeUalasnleUysasphglnly 57

RESULT 32
US-10-203-135-32559
Sequence 32559, Application US/10203135
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 5
CURRENT APPLICATION NUMBER: US/10/203,135
CURRENT FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
```

```

; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 37012
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 32559
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004848.1
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; FEATURE:
; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
US-10-203-135-32559

Alignment Scores:
Pred. No.: 2,31e-43 Length: 57
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0
DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-203-135-32559 (1-57)

QY 311 GAAATGTTGAATCCAAATGATGATCACTTCAATGGCTTGCCCAAGCTCCAGTTAT 370
DB 5 Gtmetleugluserashnsnvalletthrphenangilyleuhalasnsersersertyr 24

QY 371 CATACCTTCCTTTGGATGAGGAGCGAGTAGGCTGTATGTTGGAGCAAGATCACATA 430
DB 25 Hsthrphenleuashspgluargserargleutyrglalyalysasphstille 44

QY 431 TTTTCATTCGACCTGTTAATATCAAGATTTTCAAAAG 469
DB 45 Pheserphenaspleuvalasnllelysaaphegnllys 57

RESULT 33
US-10-203-136-33412
; Sequence 33412, Application US/10203136
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 3
; CURRENT APPLICATION NUMBER: US/10/203,136
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 38578
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
```

```

; SEQ ID NO 33412
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004848.1
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
US-10-203-136-33412

Alignment Scores:
Pred. No.: 2,31e-43 Length: 57
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0
DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-203-136-33412 (1-57)

QY 311 GAAATGTTGAATCCAAATGATGATCACTTCAATGGCTTGCCCAAGCTCCAGTTAT 370
DB 5 Gtmetleugluserashnsnvalletthrphenangilyleuhalasnsersersertyr 24

QY 371 CATACCTTCCTTTGGATGAGGAGCGAGTAGGCTGTATGTTGGAGCAAGATCACATA 430
DB 25 Hsthrphenleuashspgluargserargleutyrglalyalysasphstille 44

QY 431 TTTTCATTCGACCTGTTAATATCAAGATTTTCAAAAG 469
DB 45 Pheserphenaspleuvalasnllelysaaphegnllys 57

RESULT 34
US-10-203-137-33541
; Sequence 33541, Application US/10203137
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 7
; CURRENT APPLICATION NUMBER: US/10/203,137
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 38837
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 33541
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004848.1
; FEATURE:
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; FEATURE:
```

OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
US-10-203-33541

Alignment Scores:

Pred. No.: 2.31e-43 Length: 57
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0
DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-203-137-33541 (1-57)

QY 311 GAAATGTTGGAATCCAAATGATGATCACTTTCAATGCTGGCCAAAGCTCCAGTTAT 370
DB 5 Gtmetleugluserashnsnvallettrpneanglyleualaasnsersersertyr 24
QY 371 CATACCTTCCTTTGGATGAGAACGAGTAGGCTGTATGTTGGACCAAGATCACATA 430
DB 25 Hsthrpneleuleuspclugluargsersrtyrvalgllylalsasphistile 44
QY 431 TTTTCATTCGACCTGTTAATATCAAGATTTTCAAG 469
DB 45 PneserPneaspLeuValasnlelysaSpneGlnly 57

RESULT 35
US-10-203-139-32331

Sequence 32331, Application US/10203139

GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 4
CURRENT FILING DATE: 2002-08-02
PRIOR FILING DATE: US/10/203,139
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 37156
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 32331
LENGTH: 57
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO ACO04848.1
FEATURE:
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
FEATURE:
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
US-10-203-139-32331

Alignment Scores:

Pred. No.: 2.31e-43 Length: 57
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0

DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-203-139-32331 (1-57)

QY 311 GAAATGTTGGAATCCAAATGATGATCACTTTCAATGCTGGCCAAAGCTCCAGTTAT 370
DB 5 Gtmetleugluserashnsnvallettrpneanglyleualaasnsersersertyr 24
QY 371 CATACCTTCCTTTGGATGAGAACGAGTAGGCTGTATGTTGGACCAAGATCACATA 430
DB 25 Hsthrpneleuleuspclugluargsersrtyrvalgllylalsasphistile 44
QY 431 TTTTCATTCGACCTGTTAATATCAAGATTTTCAAG 469
DB 45 PneserPneaspLeuValasnlelysaSpneGlnly 57

RESULT 36
US-10-029-386-29372

Sequence 29372, Application US/10029386

GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
FILE REFERENCE: AEWICA-X-2
CURRENT FILING DATE: US/10/029,386
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 29372
LENGTH: 60
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR7.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.5
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.68
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.64
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.7
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.69
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
US-10-029-386-29372

Alignment Scores:

Pred. No.: 2.3e-43 Length: 60
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0
DB: 26 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-029-386-29372 (1-60)

QY 311 GAAATGTTGGAATCCAAATGATGATCACTTTCAATGCTGGCCAAAGCTCCAGTTAT 370
DB 4 Gtmetleugluserashnsnvallettrpneanglyleualaasnsersersertyr 23
QY 371 CATACCTTCCTTTGGATGAGAACGAGTAGGCTGTATGTTGGACCAAGATCACATA 430
DB 24 Hsthrpneleuleuspclugluargsersrtyrvalgllylalsasphistile 43
QY 431 TTTTCATTCGACCTGTTAATATCAAGATTTTCAAG 469
DB 44 PneserPneaspLeuValasnlelysaSpneGlnly 56

RESULT 37
US-09-500-746-13

Sequence 13, Application US/09500746
GENERAL INFORMATION:
APPLICANT: Winchester, Robert J.
APPLICANT: Gulko, Percio
APPLICANT: Seki, Tetsumori

```

/ TITLE OF INVENTION: USBS OF INHIBITORS FOR THE ACTIVATION OF CXCR4 RECEPTOR BY SDF-1
/ TITLE OF INVENTION: TREATING RHEUMATOID ARTHRITIS
/ FILE REFERENCE: 0575/57005-B
/ CURRENT APPLICATION NUMBER: US/09/500,746
/ CURRENT FILING DATE: 2000-02-09
/ NUMBER OF SEQ ID NOS: 23
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 13
/ LENGTH: 101
/ TYPE: PRT
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: MISC FEATURE
/ LOCATION: (54) (54)
/ OTHER INFORMATION: X=to any amino acid
US-09-500-746-13

Alignment Scores:
Pred. No.: 2,23e-43 Length: 101
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.92% Indels: 0
DB: 19 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-500-746-13 (1-101)

OY 1211 TCAGCCGTGTGTATGATGATGAGAGAGGCTGCTTGGTCCATATGCC 1270
Db 1 SerAlaValCymetylsermetserAspValArgValPheIuenglyProTyrAla 20

OY 1271 CACAGGATNGACCCCAACTATCAATGGGCGCTTATCAAGAGAAGATCCCCATATCCACGG 1330
Db 21 HisArgAspGlyPrcAsnTyrGlnTyrValProTyrGlnGlyArgValProTyrProArg 40

OY 1331 CCAGGAAGCTGTCCAGCAAAACATTGGTGGTTTGAC 1369
Db 41 ProGlyThrCysProSerLysThrPheGlyGlyPheAsp 53

RESULT 38
US-60-160-203-3958
/ Sequence 3958, Application US/60160203
/ GENERAL INFORMATION:
/ APPLICANT: BONAZZI, VIVIAN
/ TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
/ TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
/ TITLE OF INVENTION: USES THEREOF
/ FILE REFERENCE: CLO00116
/ CURRENT APPLICATION NUMBER: US/60/160,203
/ CURRENT FILING DATE: 1999-10-19
/ NUMBER OF SEQ ID NOS: 6374
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3958
/ LENGTH: 50
/ TYPE: PRT
/ ORGANISM: HUMAN
US-60-160-203-3958

Alignment Scores:
Pred. No.: 2,69e-38 Length: 50
Score: 48.00 Matches: 48
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.36% Indels: 0
DB: 31 Gaps: 0

US-09-774-490-1 (1-2709) x US-60-160-203-3958 (1-50)

OY 866 GATCCAAAGTTCATTAGTCCACACTCATCTCAGAGAGTGACATCTGAAGATGACAA 925
Db 3 AspProLysPheIleSerAlaHisLeuIleSerGlnSerIleAspAsnProGluAspLys 22

OY 926 GTATACCTTTCTTCGCGAATAATGATAGTGAGAAACACTCTGGAAAAGCTACTAC 985

```

```

Db      23 ValTyrPhePhePheArgGluSnaIaIleAepGlyGluHisSerGlyLysAlaThrHis 42
          |||||
          986 GCTAGAAATGCTCAGATATGCAAG 1009
          |||||
          43 AlaArgIleGlyGlnIleCysLys 50

RESULT 39
US-60-160-203-4025
; Sequence 4025, Application US/60160203
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIAN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000116
; CURRENT APPLICATION NUMBER: US/60/160,203
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 6374
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4025
; LENGTH: 50
; TYPE: PRT
; ORGANISM: HUMAN
US-60-160-203-4025

Alignment Scores:
Pred. No.:      2,69e-38      Length:      50
Score:          48.00         Matches:      48
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:    0
Query Match:    5.36%         Indels:      0
DB:             31           Gaps:          0

US-09-774-490-1 (1-2709) x US-60-160-203-4025 (1-50)

QY      866 GATCGAAAGTTCATTAGTGGCCACCTCATCTGAGAGATGCAATCCTGAAGATGACAA 925
          |||||
          Db      3 AspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAspAspLys 22
          |||||
          QY      926 GTATCTTTTCTTCCTCGTGAATAATGCATATGATGAGAACACACTGTGAAAGCTACTCAC 985
          |||||
          Db      23 ValTyrPhePhePheArgGluSnaIaIleAepGlyGluHisSerGlyLysAlaThrHis 42
          |||||
          QY      986 GCTAGAAATGCTCAGATATGCAAG 1009
          |||||
          Db      43 AlaArgIleGlyGlnIleCysLys 50

RESULT 40
US-09-864-761-44552
; Sequence 44552, Application US/09864761
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME- DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30

```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44552
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006322.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.46
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.54
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56
; OTHER INFORMATION: EST_HUMAN HIT: BF700780.1, EVALUE 1.00e-17
; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 1.00e-18
US-09-774-490-1 (1-2709) x US-09-864-761-44552 (1-41)
Alignment Scores:
Pred. No.: 3,32e-31 Length: 41
Score: 41.00 Matches: 41
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.58% Indels: 0
Gaps: 0
DB: 23
US-09-774-490-1 (1-2709) x US-09-864-761-44552 (1-41)
QY 1937 CCTGAAGAGAGATCATGTATGTTAGAGATAGTATGATTTTGGATGAGTCCG 1996
|||||
DB 1 ProgluGluArgIleIleTyGlyValGluAnsSerThrPheLeuGluCysSerPro 20
|||||
QY 1997 AAGTCGACAGAGCCGCTGTATTGCAATTCAGAGCGCAATGAAGCGAAGAAAGAA 2056
|||||
DB 21 LysSerGlnArgAlaLeuValTyrTrpGlnPheGlnArgGlnGluArgGlyGln 40
|||||
QY 2057 GAG 2059
|||||
DB 41 Gln 41
RESULT 41
US-10-182-993-33490
; Sequence 33490, Application US/10182993
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN BRAIN
; FILE REFERENCE: PB 0004 WO 2
```

```

; CURRENT APPLICATION NUMBER: US/10/182,993
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 37811
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 33490
; LENGTH: 41
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006322.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.54
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56
; OTHER INFORMATION: EST_HUMAN HIT: BF700780.1, EVALUE 1.00e-17
; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 1.00e-18
US-10-182-993-33490
Alignment Scores:
Pred. No.: 3,32e-31 Length: 41
Score: 41.00 Matches: 41
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.58% Indels: 0
Gaps: 0
DB: 27
US-09-774-490-1 (1-2709) x US-10-182-993-33490 (1-41)
QY 1937 CCTGAAGAGAGATCATGTATGTTAGAGATAGTATGATTTTGGATGAGTCCG 1996
|||||
DB 1 ProgluGluArgIleIleTyGlyValGluAnsSerThrPheLeuGluCysSerPro 20
|||||
QY 1997 AAGTCGACAGAGCCGCTGTATTGCAATTCAGAGCGCAATGAAGCGAAGAAAGAA 2056
|||||
DB 21 LysSerGlnArgAlaLeuValTyrTrpGlnPheGlnArgGlnGluArgGlyGln 40
|||||
QY 2057 GAG 2059
|||||
DB 41 Gln 41
RESULT 42
US-10-203-135-33728
; Sequence 33728, Application US/10203135
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN LUNG
; FILE REFERENCE: PB 0004 WO 5
; CURRENT APPLICATION NUMBER: US/10/203,135
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
```

```
/ PRIOR FILING DATE: 03 August 2000 (03.08.00)
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 03 October 2000 (03.10.00)
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 27 September 2000 (27.09.00)
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 21 September 2000 (21.09.00)
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 30 June 2000 (30.06.00)
/ NUMBER OF SEQ ID NOS: 37012
/ SOFTWARE: Molecular Dynamics Sequence Listing Engine
/ SEQ ID NO 33728
/ LENGTH: 41
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC006322.2
/ FEATURE:
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.54
/ FEATURE:
/ OTHER INFORMATION: EST_HUMAN HIT: BF700780.1, EVALUE 1.00e-17
/ FEATURE:
/ OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 1.00e-18
/ US-10-203-135-33728

Alignment Scores:
Pred. No.: 3,32e-31 Length: 41
Score: 41.00 Matches: 41
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.58% Indels: 0
Gaps: 0
DB:

US-09-774-490-1 (1-2709) x US-10-203-135-33728 (1-41)

QY 1937 CCTGAAGAGAGATCATCTATGCTGAGAGATGACATTTTGGATGCACTCGG 1996
DB 1 ProgluinarGileiletyrGlyValGluanserSerThrPheleugluCysserPro 20

QY 1997 AAGTCGACAGAGCGGCTGCTATTGGCAATTCAGAGCGCAATGAAGCGAAGAA 2056
DB 21 LysSerGlnArgAlaLeuValItyrTrpGlnPheGlnArgArgangluGluArgLysGlu 40

QY 2057 GAG 2059
DB 41 Glu 41

RESULT 43
US-10-203-136-34569
/ Sequence 34569, Application US/10203136
/ GENERAL INFORMATION:
/ APPLICANT: Molecular Dynamics, Inc.
/ APPLICANT: Penn. Sharron G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER
/ FILE REFERENCE: PB 0004 WO 3
/ CURRENT APPLICATION NUMBER: US/10/203,136
/ CURRENT FILING DATE: 2002-08-02
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 04 February 2000 (04.02.00)
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 26 May 2000 (26.05.00)
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 03 August 2000 (03.08.00)
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 03 October 2000 (03.10.00)
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 27 September 2000 (27.09.00)
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 21 September 2000 (21.09.00)
```

```
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 30 June 2000 (30.06.00)
/ NUMBER OF SEQ ID NOS: 38578
/ SOFTWARE: Molecular Dynamics Sequence Listing Engine
/ SEQ ID NO 34569
/ LENGTH: 41
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC006322.2
/ FEATURE:
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.46
/ FEATURE:
/ OTHER INFORMATION: EST_HUMAN HIT: BF700780.1, EVALUE 1.00e-17
/ FEATURE:
/ OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 1.00e-18
/ US-10-203-136-34569

Alignment Scores:
Pred. No.: 3,32e-31 Length: 41
Score: 41.00 Matches: 41
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.58% Indels: 0
Gaps: 0
DB:

US-09-774-490-1 (1-2709) x US-10-203-136-34569 (1-41)

QY 1937 CCTGAAGAGAGATCATCTATGCTGAGAGATGACACATTTTGGATGCACTCGG 1996
DB 1 ProgluinarGileiletyrGlyValGluanserSerThrPheleugluCysserPro 20

QY 1997 AAGTCGACAGAGCGGCTGCTATTGGCAATTCAGAGCGCAATGAAGCGAAGAA 2056
DB 21 LysSerGlnArgAlaLeuValItyrTrpGlnPheGlnArgArgangluGluArgLysGlu 40

QY 2057 GAG 2059
DB 41 Glu 41

RESULT 44
US-60-177-571-3274
/ Sequence 3274, Application US/60177571
/ GENERAL INFORMATION:
/ APPLICANT: Bonazzi, Vivien
/ TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
/ TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS,
/ FILE REFERENCE: CL000201
/ CURRENT APPLICATION NUMBER: US/60/177,571
/ CURRENT FILING DATE: 2000-01-27
/ NUMBER OF SEQ ID NOS: 5082
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3274
/ LENGTH: 53
/ TYPE: PRT
/ ORGANISM: HUMAN
/ US-60-177-571-3274

Alignment Scores:
Pred. No.: 3,55e-28 Length: 53
Score: 38.00 Matches: 38
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.25% Indels: 0
Gaps: 0
DB:

US-09-774-490-1 (1-2709) x US-60-177-571-3274 (1-53)

QY 1916 GATATTCACATGCGCAAGCCCTGAAGAGAGATCATCTATGCTGAGAGATGTAAC 1975
DB 5 AspaenHshAsglyHisSerProgluinarGileiletyrGlyValGluanserSer 24
```


QY 1976 ACATTTTGGATGCTGCTCGAAGTGGCAGAGAGCCGCTGCTATTGGCAATTC 2029
DB 25 ThrPheLeuGlnGlySerProGlySerGlnArgAlaLeuValTyrTrpGlnPhe 42

RESULT 45
US-09-500-746-9
Sequence 9, Application US/09500746
GENERAL INFORMATION:
APPLICANT: Winchester, Robert J.
APPLICANT: Gulko, Percio
APPLICANT: Seki, Tetsumori
TITLE OF INVENTION: USES OF INHIBITORS FOR THE ACTIVATION OF CXCR4 RECEPTOR BY SDF-1
FILE REFERENCE: 0575/57005-B
CURRENT APPLICATION NUMBER: US/09/500,746
NUMBER OF SEQ ID NOS: 23
SOFTWARE: PatentIn version 3.1
SEQ ID NO 9
LENGTH: 507
TYPE: PRT
ORGANISM: mouse
FEATURE:
NAME/KEY: MISC FEATURE
LOCATION: (337)..(337)
NAME/KEY: MISC FEATURE
LOCATION: (376)..(376)
OTHER INFORMATION: x= to any amino acid
US-09-500-746-9

Alignment Scores:
Pred. No.: 3,346-25 Length: 507
Score: 35.00 Matches: 35
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.91% Indels: 0
DB: 19 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-500-746-9 (1-507)

QY 1241 GTGAGAGGGTGTCTCTGTCCTATATGCGCCACAGGATGACCCCACTATCAATGAGGTG 1300
DB 174 ValArgGArgValPheLeuGlnGlyProTyrAlaHisArgGArgGlyProAsnTyrGlnTrpVal 193

QY 1301 CCTTTCAGAGAGAGTCCCTATCCAGCGCCAGCAACTTGTGCC 1345
DB 194 ProTyrGlnGlyArgValProTyrProArgProGlyThrCysPro 208

RESULT 46
PCT-US01-00663-33648
Sequence 33648, Application PC/TUS0100663
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wenheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 7
CURRENT APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)

PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38837
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 33648
LENGTH: 33
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUATION 4.00e-15
PCT-US01-00663-33648

Alignment Scores:
Pred. No.: 4,226-23 Length: 33
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US01-00663-33648 (1-33)

QY 533 AAGATGCTCAATTGATCAAGTACTTAAGCATTAATGACACTGTTACGCC 592
DB 1 LysGlnCysAlaAsnPheLeuValLeuValAlaTyrAsnGlnThrHisLeuTyrAla 20

QY 593 TGTGAGCGGGGCTTTTCATCCATTTGCACTTACATT 631
DB 21 CysGlnThrGlyAlaPheHisProLeuIleCysThrTyrIle 33

RESULT 47
US-09-864-761-39708
Sequence 39708, Application US/09864761
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wenheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 39708
LENGTH: 33
TYPE: PR1
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN PETAL LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 4.00e-15
US-09-761-39708

Alignment Scores:
Pred. No.: 4.22e-23 Length: 33
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 23 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-864-761-39708 (1-33)

QY 533 AAGATGTCATTTTCATCAAGTACTTAAGCATATATACAGACTGTTAGGCC 592
|||||
DB 1 LysGluCyAlaAsnPhelIelysValLeuYalaIryAsnInThHisLeuYzala 20

QY 593 TGTGAACGGGGCTTTTCATTCACATTTGCACCTACATT 631
|||||
DB 21 CysGlyThrGlyAlaPheHisProIleCysThrTyIle 33

RESULT 48
US-10-182-993-32618
Sequence 32618, Application US/10182993
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn. Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 2
CURRENT APPLICATION NUMBER: US/10/182,993
PRIOR FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 37811
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 32618
LENGTH: 33
TYPE: PR1
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 4.00e-15
US-10-182-993-32618

Alignment Scores:
Pred. No.: 4.22e-23 Length: 33
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 27 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-182-993-32618 (1-33)

QY 533 AAGATGTCATTTTCATCAAGTACTTAAGCATATATACAGACTGTTAGGCC 592
|||||
DB 1 LysGluCyAlaAsnPhelIelysValLeuYalaIryAsnInThHisLeuYzala 20

QY 593 TGTGAACGGGGCTTTTCATTCACATTTGCACCTACATT 631
|||||
DB 21 CysGlyThrGlyAlaPheHisProIleCysThrTyIle 33

RESULT 49
US-10-182-995-26180
Sequence 26180, Application US/10182995
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn. Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 1
CURRENT APPLICATION NUMBER: US/10/182,995
PRIOR FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
NUMBER OF SEQ ID NOS: 29119
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 26180
LENGTH: 33
TYPE: PR1
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 4.00e-15

US-10-182-995-26180

Alignment Scores:

Pred. No.:	4,22e-23	Length:	33
Score:	33.00	Matches:	33
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.69%	Indels:	0
DB:	27	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-182-995-26180 (1-33)

QY 533 AAAGATGCTAATTTCATCAAGGTAAGGATATATACAGACTGTTGACGCC 592

DB 1 Lysglucyalaasphenlelevalleuysalaetyraenglnthrhisleutyrala 20

QY 593 TGTGAACGGGGCTTTTCATTCACATTTGCACCTTACATT 631

DB 21 Cysglythnglyalaphheisprolelecystrnyrile 33

RESULT 50

US-10-203-134-33472

; Sequence 33472, Application US/10203134

; GENERAL INFORMATION:

; APPLICANT: Molecular Dynamics, Inc.

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: PB 0004 WO 6

; CURRENT FILING DATE: 2002-08-02

; PRIOR FILING DATE: 2002-08-02

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 04 February 2000 (04.02.00)

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 26 May 2000 (26.05.00)

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 03 August 2000 (03.08.00)

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 03 October 2000 (03.10.00)

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 27 September 2000 (27.09.00)

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 21 September 2000 (21.09.00)

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 30 June 2000 (30.06.00)

; NUMBER OF SEQ ID NOS: 38628

; SOFTWARE: Molecular Dynamics Sequence Listing Engine

; SEQ ID NO 33472

; LENGTH: 33

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC004848.1

; FEATURE:

; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6

; FEATURE:

; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUATION = 4.00e-15

; US-10-203-134-33472

Alignment Scores:

Pred. No.:	4,22e-23	Length:	33
Score:	33.00	Matches:	33
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.69%	Indels:	0
DB:	28	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-203-134-33472 (1-33)

QY 533 AAAGATGCTAATTTCATCAAGGTAAGGATATATACAGACTGTTGACGCC 592

DB 1 Lysglucyalaasphenlelevalleuysalaetyraenglnthrhisleutyrala 20

QY 593 TGTGAACGGGGCTTTTCATTCACATTTGCACCTTACATT 631

DB 21 Cysglythnglyalaphheisprolelecystrnyrile 33

RESULT 51

US-10-203-135-32678

; Sequence 32678, Application US/10203135

; GENERAL INFORMATION:

; APPLICANT: Molecular Dynamics, Inc.

; APPLICANT: Penn, Sharron G.

; APPLICANT: Rank, David R.

; APPLICANT: Hanzel, David K.

; APPLICANT: Chen, Wensheng

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

; FILE REFERENCE: PB 0004 WO 5

; CURRENT FILING DATE: 2002-08-02

; PRIOR FILING DATE: 2002-08-02

; PRIOR APPLICATION NUMBER: US 60/180,312

; PRIOR FILING DATE: 04 February 2000 (04.02.00)

; PRIOR APPLICATION NUMBER: US 60/207,456

; PRIOR FILING DATE: 26 May 2000 (26.05.00)

; PRIOR APPLICATION NUMBER: US 09/632,366

; PRIOR FILING DATE: 03 August 2000 (03.08.00)

; PRIOR APPLICATION NUMBER: GB 24263.6

; PRIOR FILING DATE: 03 October 2000 (03.10.00)

; PRIOR APPLICATION NUMBER: US 60/236,359

; PRIOR FILING DATE: 27 September 2000 (27.09.00)

; PRIOR APPLICATION NUMBER: US 60/234,687

; PRIOR FILING DATE: 21 September 2000 (21.09.00)

; PRIOR APPLICATION NUMBER: US 09/608,408

; PRIOR FILING DATE: 30 June 2000 (30.06.00)

; NUMBER OF SEQ ID NOS: 37012

; SOFTWARE: Molecular Dynamics Sequence Listing Engine

; SEQ ID NO 32678

; LENGTH: 33

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: MAP TO AC004848.1

; FEATURE:

; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5

; FEATURE:

; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUATION = 4.00e-15

; US-10-203-135-32678

Alignment Scores:

Pred. No.:	4,22e-23	Length:	33
Score:	33.00	Matches:	33
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	3.69%	Indels:	0
DB:	28	Gaps:	0

US-09-774-490-1 (1-2709) x US-10-203-135-32678 (1-33)

QY 533 AAAGATGCTAATTTCATCAAGGTAAGGATATATACAGACTGTTGACGCC 592

DB 1 Lysglucyalaasphenlelevalleuysalaetyraenglnthrhisleutyrala 20

QY 593 TGTGAACGGGGCTTTTCATTCACATTTGCACCTTACATT 631

DB 21 Cysglythnglyalaphheisprolelecystrnyrile 33

RESULT 52

US-10-203-136-33531

; Sequence 33531, Application US/10203136

; GENERAL INFORMATION:

; APPLICANT: Molecular Dynamics, Inc.

; APPLICANT: Penn, Sharron G.

```
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 3
CURRENT APPLICATION NUMBER: US/10/203,136
PRIOR FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38578
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 33531
LENGTH: 33
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
FEATURE:
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 4.00e-15
US-10-203-136-33531

Alignment Scores:
Pred. No.: 4,22e-23 Length: 33
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-203-136-33531 (1-33)
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 7
CURRENT APPLICATION NUMBER: US/10/203,137
PRIOR FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
US-10-203-137-33648
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN PLACENTA
FILE REFERENCE: PB 0004 WO 7
CURRENT APPLICATION NUMBER: US/10/203,137
PRIOR FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
```

```
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38837
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 33648
LENGTH: 33
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
FEATURE:
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 4.00e-15
US-10-203-137-33648

Alignment Scores:
Pred. No.: 4,22e-23 Length: 33
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.69% Indels: 0
DB: 28 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-203-137-33648 (1-33)
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 4
CURRENT APPLICATION NUMBER: US/10/203,139
PRIOR FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 37156
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 32427
LENGTH: 33

593 TGTGAACGGGGCTTTTCATCAGGCTTAAGGCTATATACAGCTGTTACGCC 592
1 LysGluCyAlaAsnPhelIeYsValIeuYsAlaTyAsnGlnThrhIstLeuYrAla 20
21 CysGlyThnGlyAlaPhelIstProIleCysThrTyIle 33

593 TGTGAACGGGGCTTTTCATCAGGCTTAAGGCTATATACAGCTGTTACGCC 631
1 LysGluCyAlaAsnPhelIeYsValIeuYsAlaTyAsnGlnThrhIstLeuYrAla 20
21 CysGlyThnGlyAlaPhelIstProIleCysThrTyIle 33

593 TGTGAACGGGGCTTTTCATCAGGCTTAAGGCTATATACAGCTGTTACGCC 631
1 LysGluCyAlaAsnPhelIeYsValIeuYsAlaTyAsnGlnThrhIstLeuYrAla 20
21 CysGlyThnGlyAlaPhelIstProIleCysThrTyIle 33

593 TGTGAACGGGGCTTTTCATCAGGCTTAAGGCTATATACAGCTGTTACGCC 631
1 LysGluCyAlaAsnPhelIeYsValIeuYsAlaTyAsnGlnThrhIstLeuYrAla 20
21 CysGlyThnGlyAlaPhelIstProIleCysThrTyIle 33
```



```

: TITLE OF INVENTION: US THEREOF
: FILE REFERENCE: CLO00137
: CURRENT APPLICATION NUMBER: US/60/163,123
: CURRENT FILING DATE: 1999-11-02
: NUMBER OF SEQ ID NOS: 1986
: SOFTWARE: PatsSeq for Windows Version 4.0
: SEQ ID NO 1597
: LENGTH: 35
: TYPE: PRT
: ORGANISM: Human
US-60-163-123-1597

```

Alignment Scores:	
Pred. NO.:	4.45e-21
Score:	31.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	3.46%
OB:	31
Length:	35
Matches:	31
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-774-490-1 (1-2709) X US-60-163-123-1597 (1-35)

Qy 1555 GATGTGGGACCGTCTTAAAGTGTTCATTCCTAAGAGACTTGGATGATTAGAA 1618
Db 1 AspValGlyThrvalLeuysvalValSerIlepolysGlnumTrpPyraSpenglu 20

```

QY      1619 GAGGTTCTGCTGGAAAGATGACACTTTTTCGG 1651
          |||||
Db      21  GluValLeuLeuGluGluMetThrValPheArg 31

```

RESULT 59
US-09-791-537-105918
; Sequence 105918, Application US/09791537

```

?
?
? GENERAL INFORMATION:
? APPLICANT: Biomimix, Inc.
? APPLICANT: Debe, Derek
? APPLICANT: Danzer, Joseph
? TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
? FILE OF INVENTION: METHODS OF USE THEREOF
? FILE REFERENCE: 261/210
? CURRENT APPLICATION NUMBER: US/09/791,537
? CURRENT FILING DATE: 2001-02-22
? NUMBER OF SEQ ID NOS: 153055
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 105918
?

```

```

; TYPE: PRT
; ORGANISM: Danio rerio
US-09-791-537-105918

```

Alignment Scores:	
Pred. No.:	4,44e-14
Score:	24.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	2.68%
BB:	22
Length:	778
Matches:	24
Conservative:	0
Mismatches:	0
Indels:	0
Gaps:	0

US-09-774-490-1 (1-2709) x US-09-791-537-105918 (1-778)

Qy 123 TGTTCCTGCTATTTTCCACTGCAAGAAGACGCAAGAAGCAAGAATTAAGAATGGA 1882

Dd 543 CysSerArgTyrPheProThrAlaAlaArgThrThrArgGlnAspLeuArgAsnGly 562

QY	1883	GACCCACTGACT	1894
Db	563	AspProLeuThr	566

RESULT 60
US-09-513-999C-6632
; Sequence 6632, Application US/09513999C
; GENERAL INFORMATION:
; APPLICANT: Dumas Mline Edwards, J.B.

```

: APPLICANT: Ducleert, A.
: APPLICANT: Giordano, J.Y.
: TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins
: FILE REFERENCE: 59,US2.REG
: CURRENT APPLICATION NUMBER: US/09/513,999C
: CURRENT FILING DATE: 2000-02-24
: PRIOR APPLICATION NUMBER: US 60/122,487
: PRIOR FILING DATE: 1999-02-26
: NUMBER OF SEQ ID NOS: 36681
: SOFTWARE: Patent.pm
: SEQ ID NO 6632

```

```

? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: UNSURE
? LOCATION: 28
? OTHER INFORMATION: Xaa=Glu or Lys
? FEATURE:
? NAME/KEY: UNSURE
? LOCATION: 29

```

```

; OTHER INFORMATION: Xaa=Ala or Thr
;
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 30
; OVER INFORMATION:

```

```
;
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 32
```

US-09-513-999C-6632

Alignment Scores:

Pred. No.:	5,51e-12	length:	67
Score:	22.00	Matches:	35
Percent Similarity:	97.22%	Conservative:	0
Best Local Similarity:	97.22%	Mismatches:	0
Query Match:	2.46%	Indels:	1
DB:	19	Gaps:	0

US-09-774-490-1 (1-2709) x US-09-513-999C-6632 (1-67)

176 AAAGGACCTACAGCGTCTCCAGCATGGGCTGGTTAACTAGAGATTGTCTGTCTTTTCTGG 235
 33 LysGlyIhrTyrSerValCysSerMetGlyTrpLeuThrArgIleValCysLeuPheTrp 52

OY 236 GGAGTATTACTTACAGCAAGCAAACTATCAGAATGGGAGAAC 281
| | | | | | | | | | | | | | | | | | | | | |
DB 53 -glutyrtyrleuglnglntthrileargmetglyargthr 67

RESULT 61
PCT-US00-

```
; Sequence 16, Application PC/TUS0041943
; GENERAL INFORMATION:
```

; TITLE OF INVENTION: NOVEL SEMAPHORIN DOMAINS

```

; FILE REFERENCE: 00786-441W01
; CURRENT APPLICATION NUMBER: PCT/US00/41943

```

; CURRENT FILING DATE: 2000-11-07
 ; PRIOR APPLICATION NUMBER: US 09/694,085

! PRIOR FILING DATE: 2000-10-20
! PRIOR APPLICATION NUMBER: US 60/164,056

```

; PRIOR FILING DATE: 1999-11-08
; NUMBER OF SEQ ID NOS: 48

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16

```

```

; LENGTH: 17
; TYPE: PRT
;

```

```

; ORGANISM: Artificial Sequence
; FEATURE:

```

OTHER INFORMATION: computer-generated amino acid
PCT-US00-41943-16

Alignment Scores:
Pred. No.: 6.93e-07 Length: 17
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US00-41943-16 (1-17)

QY 1763 AAGCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTTGGAT 1813

DB 1 LysAlaCyAlaGluCysCysLeuAlaArgAspProTyrCysAlaItrpAsp 17

RESULT 62
PCT-US00-41943-43

/ Sequence 43, Application PC/TUS0041943
/ GENERAL INFORMATION:
/ APPLICANT: The General Hospital Corporation
/ TITLE OF INVENTION: NOVEL SEMAPHORIN DOMAINS
/ FILE REFERENCE: 00786-441W01
/ CURRENT APPLICATION NUMBER: PCT/US00/41943
/ PRIOR FILING DATE: 2000-11-07
/ PRIOR APPLICATION NUMBER: US 09/694,085
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: US 60/164,056
/ PRIOR FILING DATE: 1999-11-08
/ NUMBER OF SEQ ID NOS: 48
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 43
/ LENGTH: 17
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: computer-generated amino acid
PCT-US00-41943-43

Alignment Scores:
Pred. No.: 6.93e-07 Length: 17
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US00-41943-43 (1-17)

QY 1763 AAGCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTTGGAT 1813

DB 1 LysAlaCyAlaGluCysCysLeuAlaArgAspProTyrCysAlaItrpAsp 17

RESULT 63

US-09-694-085-16
/ Sequence 16, Application US/09694085
/ GENERAL INFORMATION:
/ APPLICANT: Behar, Oded
/ TITLE OF INVENTION: NOVEL SEMAPHORIN DOMAINS
/ FILE REFERENCE: 00786-441W01
/ CURRENT APPLICATION NUMBER: US/09/694,085
/ CURRENT FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: US 60/164,056
/ PRIOR FILING DATE: 1999-11-08
/ NUMBER OF SEQ ID NOS: 48
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 16
/ LENGTH: 17
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: computer-generated amino acid
US-09-694-085-16

Alignment Scores:
Pred. No.: 6.93e-07 Length: 17
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 20 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-694-085-16 (1-17)

QY 1763 AAGCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTTGGAT 1813

DB 1 LysAlaCyAlaGluCysCysLeuAlaArgAspProTyrCysAlaItrpAsp 17

RESULT 64
US-09-694-085-43

/ Sequence 43, Application US/09694085
/ GENERAL INFORMATION:
/ APPLICANT: Behar, Oded
/ TITLE OF INVENTION: NOVEL SEMAPHORIN DOMAINS
/ FILE REFERENCE: 00786-441W01
/ CURRENT APPLICATION NUMBER: US/09/694,085
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: US 60/164,056
/ PRIOR FILING DATE: 1999-11-08
/ NUMBER OF SEQ ID NOS: 48
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 43
/ LENGTH: 17
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: computer-generated amino acid
US-09-694-085-43

Alignment Scores:
Pred. No.: 6.93e-07 Length: 17
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 20 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-694-085-43 (1-17)

QY 1763 AAGCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTTGGAT 1813

DB 1 LysAlaCyAlaGluCysCysLeuAlaArgAspProTyrCysAlaItrpAsp 17

RESULT 65

US-60-177-571-3765
/ Sequence 3765, Application US/60177571
/ GENERAL INFORMATION:
/ APPLICANT: Bonazzi, Vivien
/ TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
/ TITLE OF INVENTION: NOCTERIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS,
/ FILE REFERENCE: CL000201
/ CURRENT APPLICATION NUMBER: US/60/177,571
/ CURRENT FILING DATE: 2000-01-27
/ NUMBER OF SEQ ID NOS: 5082
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3765
/ LENGTH: 73
/ TYPE: PRT
/ ORGANISM: HUMAN
US-60-177-571-3765
Alignment Scores:
Pred. No.: 6.31e-07 Length: 73
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 31 Gaps: 0

US-09-774-490-1 (1-2709) x US-60-177-571-3765 (1-73)

QY 1766 GCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTGGATGCT 1816

DB 28 AlAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 44

RESULT 66

US-60-177-646-3105
; Sequence 3105, Application US/60177646
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,
; FILE REFERENCE: CL000210
; CURRENT APPLICATION NUMBER: US/60/177,646
; CURRENT FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 4226
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3105
; LENGTH: 73
; TYPE: PRT
; ORGANISM: HUMAN
US-60-177-646-3105

Alignment Scores:

Pred. No.: 6.31e-07 Length: 73
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 31 Gaps: 0

US-09-774-490-1 (1-2709) x US-60-177-646-3105 (1-73)

QY 1766 GCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTGGATGCT 1816

DB 28 AlAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 44

RESULT 67

US-60-177-571-3766
; Sequence 3766, Application US/60177571
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS,
; FILE REFERENCE: CL000201
; CURRENT APPLICATION NUMBER: US/60/177,571
; CURRENT FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 5082
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3766
; LENGTH: 79
; TYPE: PRT
; ORGANISM: HUMAN
US-60-177-571-3766

Alignment Scores:

Pred. No.: 6.28e-07 Length: 79
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 31 Gaps: 0

US-09-774-490-1 (1-2709) x US-60-177-571-3766 (1-79)

QY 1766 GCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTGGATGCT 1816

DB 26 AlAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 42

RESULT 68

PCT-US00-33116-2
; Sequence 2, Application PC/TUS0033116
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: HUMAN SEMAPHORIN ZSMF-16
; FILE REFERENCE: 99-90PC
; CURRENT APPLICATION NUMBER: PCT/US00/33116
; CURRENT FILING DATE: 2000-12-06
; PRIOR APPLICATION NUMBER: US 90/455,560
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 779
; TYPE: PRT
; ORGANISM: Homo Sapiens
PCT-US00-33116-2

Alignment Scores:

Pred. No.: 5.42e-07 Length: 779
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US00-33116-2 (1-779)

QY 1766 GCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTGGATGCT 1816

DB 525 AlAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 541

RESULT 69

US-09-455-560-2
; Sequence 2, Application US/09455560
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L.
; TITLE OF INVENTION: HUMAN SEMAPHORIN ZSMF-16
; FILE REFERENCE: 99-90X
; CURRENT APPLICATION NUMBER: US/09/455,560
; CURRENT FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 779
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-455-560-2

Alignment Scores:

Pred. No.: 5.42e-07 Length: 779
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 18 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-455-560-2 (1-779)

QY 1766 GCGTGTGCTGAGTGTTCCTCGCCGAGACCTTACTGTGCTGGATGCT 1816

DB 525 AlAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaTrpAspGly 541

RESULT 70

US-09-731-179-2
; Sequence 2, Application US/09731179
; GENERAL INFORMATION:
; APPLICANT: Holloway, James L..


```
/ APPLICANT: Foley, Kevin P.
/ TITLE OF INVENTION: HUMAN SEMAPHORIN ZSNF-16
/ FILE REFERENCE: 99-90
/ CURRENT APPLICATION NUMBER: US/09/731,179
/ CURRENT FILING DATE: 2000-12-06
/ PRIOR APPLICATION NUMBER: US 60/169,238
/ PRIOR FILING DATE: 1999-12-06
/ NUMBER OF SEQ ID NOS: 6
/ SOFTWARE: FastSeq for Windows Version 3.0
/ SEQ ID NO 2
/ LENGTH: 779
/ TYPE: PRT
/ ORGANISM: Homo Sapiens
US-09-731-179-2

Alignment Scores:
Pred. No.: 5,41e-07 Length: 779
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 21 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-731-179-2 (1-779)
QY 1766 GCGTGTGCTGAGTGTGCTCGCCGAGACCTTACTGTGCTGGATGAT 1816
DB 528 AAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaItrpAspGly 544

RESULT 71
PCT-US00-04340-18
/ Sequence 18, Application PC/TUS0004340
/ GENERAL INFORMATION:
/ APPLICANT: Valenzuela, Dario
/ APPLICANT: Yuan, Olive
/ APPLICANT: Hoffman, Heidi
/ APPLICANT: Hall, Jeff
/ APPLICANT: Radtke, Peter
/ TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
/ FILE REFERENCE: GI 6918X
/ CURRENT APPLICATION NUMBER: PCT/US00/04340
/ CURRENT FILING DATE: 2000-02-18
/ NUMBER OF SEQ ID NOS: 101
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 18
/ LENGTH: 782
/ TYPE: PRT
/ ORGANISM: Homo sapiens
PCT-US00-04340-18

Alignment Scores:
Pred. No.: 5,41e-07 Length: 782
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US00-04340-18 (1-782)
QY 1766 GCGTGTGCTGAGTGTGCTCGCCGAGACCTTACTGTGCTGGATGAT 1816
DB 528 AAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaItrpAspGly 544

RESULT 72
US-09-507-209-18
/ Sequence 18, Application US/09507209
/ GENERAL INFORMATION:
/ APPLICANT: Valenzuela, Dario
/ APPLICANT: Yuan, Olive
/ APPLICANT: Hoffman, Heidi
/ APPLICANT: Hall, Jeff
/ APPLICANT: Radtke, Peter
```

```
/ TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
/ FILE REFERENCE: GI 6918X
/ CURRENT APPLICATION NUMBER: US/09/507,209
/ CURRENT FILING DATE: 2000-02-18
/ NUMBER OF SEQ ID NOS: 101
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 18
/ LENGTH: 782
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-507-209-18

Alignment Scores:
Pred. No.: 5,41e-07 Length: 782
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 19 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-507-209-18 (1-782)
QY 1766 GCGTGTGCTGAGTGTGCTCGCCGAGACCTTACTGTGCTGGATGAT 1816
DB 528 AAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaItrpAspGly 544

RESULT 73
US-09-813-290-4
/ Sequence 4, Application US/09813290
/ GENERAL INFORMATION:
/ APPLICANT: Walke, D. Wade
/ APPLICANT: Wilganowski, Nathaniel L.
/ APPLICANT: Turner, C. Alexander Jr.
/ APPLICANT: Hilbun, Erin
/ APPLICANT: Wang, Xiaoming
/ APPLICANT: Donoho, Gregory
/ APPLICANT: Scoville, John
/ TITLE OF INVENTION: Novel Human Secreted Proteins and Polynucleotides Encoding the S
/ FILE REFERENCE: LEX-0151-USA
/ CURRENT APPLICATION NUMBER: US/09/813,290
/ CURRENT FILING DATE: 2001-03-20
/ PRIOR APPLICATION NUMBER: US 60/190,638
/ PRIOR FILING DATE: 2000-03-20
/ PRIOR APPLICATION NUMBER: US 60/191,188
/ PRIOR FILING DATE: 2000-03-22
/ PRIOR APPLICATION NUMBER: US 60/193,639
/ PRIOR FILING DATE: 2000-03-31
/ NUMBER OF SEQ ID NOS: 10
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 4
/ LENGTH: 782
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-813-290-4

Alignment Scores:
Pred. No.: 5,41e-07 Length: 782
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 23 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-813-290-4 (1-782)
QY 1766 GCGTGTGCTGAGTGTGCTCGCCGAGACCTTACTGTGCTGGATGAT 1816
DB 528 AAcysAlaGluCysCysLeuAlaArgAspProTyrCysAlaItrpAspGly 544

RESULT 74
PCT-US00-04340-92
/ Sequence 92, Application PC/TUS0004340
/ GENERAL INFORMATION:
```

APPLICANT: Valenzuela, Dario
APPLICANT: Yuan, Olive
APPLICANT: Hoffman, Heidi
APPLICANT: Hall, Jeff
APPLICANT: Rajceiko, Peter
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: GI 6918X
CURRENT APPLICATION NUMBER: PCT/US00/04340
CURRENT FILING DATE: 2000-02-18
NUMBER OF SEQ ID NOS: 101
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 92
LENGTH: 785
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US00-04340-92

Alignment Scores:
Pred. No.: 5,41e-07 Length: 785
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US00-04340-92 (1-785)

QY 1766 GCCTGCTGAGTGTGCTGCCCGAGACCTTACTGTGCGATGGT 1816
DB 531 AAcysAlaGluCysCysLeuAlaArgAspProtyrCysAlaTrpAspGly 547

RESULT 75

US-09-507-209-92
Sequence 92, Application US/09507209
GENERAL INFORMATION:
APPLICANT: Valenzuela, Dario
APPLICANT: Yuan, Olive
APPLICANT: Hoffman, Heidi
APPLICANT: Hall, Jeff
APPLICANT: Rajceiko, Peter
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
FILE REFERENCE: GI 6918X
CURRENT APPLICATION NUMBER: US/09/507,209
CURRENT FILING DATE: 2000-02-18
NUMBER OF SEQ ID NOS: 101
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 92
LENGTH: 785
TYPE: PRT
ORGANISM: Homo sapiens
US-09-507-209-92

Alignment Scores:
Pred. No.: 5,41e-07 Length: 785
Score: 17.00 Matches: 17
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.90% Indels: 0
DB: 19 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-507-209-92 (1-785)

QY 1766 GCCTGCTGAGTGTGCTGCCCGAGACCTTACTGTGCGATGGT 1816
DB 531 AAcysAlaGluCysCysLeuAlaArgAspProtyrCysAlaTrpAspGly 547

Search completed: August 3, 2003, 10:47:06
Job time : 459 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: August 3, 2003, 10:34:54 ; Search time 10 Seconds
(without alignments)
6791.454 Million cell updates/sec

Title: US-09-774-490-1
Perfect score: 895
Sequence: 1 aatcttattatcatgcatg.....agccttttcttcaataacc 2709

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 67526 segs, 12534983 residues

Word size: 1

Total number of hits satisfying chosen parameters: 128650

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODE=frame+.n2p.model -DEV=xlh
-Q=/cgn2_1/USFTO_spool/US09774490/runat_03082003_102929_4545/app_query.fasta_1.2887
-DB=pendling_patents_AA_New -OPMT=faetan -SUFFIX=oligo.rapn -MINMATCH=0.1
-LOOPEL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-FRAMES=human40.cdi -LIST=100 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1
-ALIGN=75 -MODE=LOCAL -OUTFRM=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USRR=US09774490 @CCN 1.1.9 @runat_03082003_102929_4545
-NCPUR=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FCGAPOP=6 -FCGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

1: Pending Patents AA_New.*
2: /cgn2_6/prodata/2/paa/PCT_NEW_COMB.pep.*
3: /cgn2_6/prodata/2/paa/US06_NEW_COMB.pep.*
4: /cgn2_6/prodata/2/paa/US07_NEW_COMB.pep.*
5: /cgn2_6/prodata/2/paa/US08_NEW_COMB.pep.*
6: /cgn2_6/prodata/2/paa/US09_NEW_COMB.pep.*
7: /cgn2_6/prodata/2/paa/US10_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	15	1.7	514	6	US-10-357-820-28	Sequence 28, Appl
2	15	1.7	636	6	US-10-408-765A-2003	Sequence 2003, Ap
3	13	1.5	775	1	PCT-US03-09929-92	Sequence 92, Appl
4	13	1.5	775	1	PCT-US03-09929-94	Sequence 94, Appl
5	13	1.5	775	1	PCT-US03-09929-96	Sequence 96, Appl
6	13	1.5	775	1	PCT-US03-09929-98	Sequence 98, Appl
7	11	1.2	431	6	US-10-369-072-97	Sequence 97, Appl
8	10	1.1	495	1	PCT-US03-09929-4	Sequence 4, Appl
9	10	1.1	634	1	PCT-US03-09929-8	Sequence 2, Appl
10	10	1.1	638	1	PCT-US03-09929-8	Sequence 8, Appl
11	10	1.1	638	1	PCT-US03-09929-10	Sequence 10, Appl

12	10	1.1	833	1	PCT-US03-09929-6	Sequence 6, Appl
13	10	1.1	833	1	PCT-US03-09929-12	Sequence 12, Appl
14	10	1.1	833	1	US-10-369-072-14	Sequence 14, Appl
15	10	1.1	833	6	US-10-369-072-16	Sequence 16, Appl
16	10	1.1	833	6	US-10-369-072-18	Sequence 18, Appl
17	10	1.1	834	6	US-10-369-072-47	Sequence 47, Appl
18	10	1.1	963	6	US-10-369-072-46	Sequence 46, Appl
19	10	1.1	350	6	US-10-376-397A-42	Sequence 42, Appl
20	9	1.0	510	6	US-10-369-072-48	Sequence 48, Appl
21	9	1.0	782	6	US-10-369-072-39	Sequence 39, Appl
22	9	1.0	893	6	US-10-369-072-38	Sequence 38, Appl
23	9	1.0	1244	6	US-10-326-956-1715	Sequence 1715, Ap
24	9	1.0	3647	6	US-10-408-765A-2436	Sequence 2436, Ap
25	8	0.9	312	7	US-60-487-610-1445	Sequence 1645, Ap
26	8	0.9	275	6	US-10-292-798-748	Sequence 748, Appl
27	8	0.9	754	6	US-10-369-072-10	Sequence 10, Appl
28	8	0.9	766	6	US-10-369-072-8	Sequence 8, Appl
29	8	0.9	760	6	US-10-369-072-36	Sequence 36, Appl
30	8	0.9	761	6	US-10-293-244-3947	Sequence 3947, Ap
31	8	0.9	762	6	US-10-369-072-35	Sequence 35, Appl
32	8	0.9	838	6	US-10-369-072-49	Sequence 49, Appl
33	8	0.9	1074	6	US-10-273-573-9338	Sequence 9338, Ap
34	8	0.9	1074	6	US-10-369-072-41	Sequence 41, Appl
35	8	0.9	1074	6	US-10-369-072-44	Sequence 44, Appl
36	8	0.9	1077	6	US-10-369-072-43	Sequence 43, Appl
37	8	0.9	5032	6	US-10-408-765A-26	Sequence 26, Appl
38	7	0.8	10	6	US-10-462-850-2385	Sequence 2385, Ap
39	7	0.8	10	6	US-10-462-850-2387	Sequence 2387, Ap
40	7	0.8	10	6	US-10-462-850-2399	Sequence 2399, Ap
41	7	0.8	10	6	US-10-462-850-2401	Sequence 2401, Ap
42	7	0.8	10	6	US-10-462-850-2485	Sequence 2485, Ap
43	7	0.8	10	6	US-10-462-850-2527	Sequence 2527, Ap
44	7	0.8	10	6	US-10-462-850-2533	Sequence 2533, Ap
45	7	0.8	10	6	US-10-462-850-2557	Sequence 2557, Ap
46	7	0.8	10	6	US-10-462-850-2557	Sequence 2557, Ap
47	7	0.8	12	6	PCT-US02-39397-69	Sequence 156, App
48	7	0.8	22	1	PCT-US02-39397-69	Sequence 156, App
49	7	0.8	47	6	US-10-369-072-90	Sequence 90, Appl
50	7	0.8	72	6	US-10-603-114-7994	Sequence 7994, Appl
51	7	0.8	121	6	US-10-374-780A-2152	Sequence 2152, Ap
52	7	0.8	126	6	US-10-273-573-7304	Sequence 1234, Ap
53	7	0.8	145	6	US-10-374-780A-1248	Sequence 1248, Ap
54	7	0.8	151	6	US-10-374-780A-1342	Sequence 1342, Ap
55	7	0.8	192	6	US-10-603-114-7828	Sequence 7828, Ap
56	7	0.8	193	6	US-10-603-114-6794	Sequence 6794, Ap
57	7	0.8	194	6	US-10-293-244-3807	Sequence 3807, Ap
58	7	0.8	231	6	US-10-374-780A-414	Sequence 414, App
59	7	0.8	231	6	US-10-374-780A-1246	Sequence 1246, Ap
60	7	0.8	251	6	US-10-603-114-4250	Sequence 4250, Ap
61	7	0.8	258	6	US-10-293-244-3649	Sequence 3649, Ap
62	7	0.8	269	6	US-10-603-114-5243	Sequence 5243, Ap
63	7	0.8	271	6	US-10-292-798-1056	Sequence 1056, Ap
64	7	0.8	287	6	US-10-603-114-5282	Sequence 5282, Ap
65	7	0.8	295	5	US-09-665-308E-2	Sequence 2, Appl
66	7	0.8	303	6	US-10-374-780A-1658	Sequence 1658, Ap
67	7	0.8	325	6	US-10-610-917-5	Sequence 5, Appl
68	7	0.8	329	6	US-10-292-798-798	Sequence 798, App
69	7	0.8	332	6	US-10-293-244-1681	Sequence 1681, Ap
70	7	0.8	333	1	PCT-US03-21576-38	Sequence 38, Appl
71	7	0.8	341	1	PCT-US03-21576-39	Sequence 39, Appl
72	7	0.8	341	1	PCT-US03-21703-13	Sequence 13, Appl
73	7	0.8	345	6	US-10-387-629-106	Sequence 106, App
74	7	0.8	351	6	US-10-374-780A-124	Sequence 124, App
75	7	0.8	352	6	US-10-341-434-178	Sequence 178, App
76	7	0.8	354	4	US-60-487-369-6	Sequence 6, Appl
77	7	0.8	355	6	US-10-403-902A-22	Sequence 22, Appl
78	7	0.8	357	6	US-10-293-244-1353	Sequence 1353, Ap
79	7	0.8	359	6	US-10-603-114-6082	Sequence 6082, Ap
80	7	0.8	360	6	US-10-374-780A-2376	Sequence 2376, Ap
81	7	0.8	370	7	US-60-487-610-1928	Sequence 1928, Ap
82	7	0.8	375	6	US-10-408-765A-1649	Sequence 1649, Ap
83	7	0.8	384	1	PCT-US03-15466-4	Sequence 4, Appl
84	7	0.8	385	6	US-10-603-114-5737	Sequence 5737, Ap

```
85 7 0.8 411 6 US-10-408-765A-126 Sequence 126, App
86 7 0.8 420 6 US-10-273-573-6661 Sequence 6661, Ap
87 7 0.8 425 6 US-09-721-870-109 Sequence 109, App
88 7 0.8 430 6 US-10-369-072-91 Sequence 91, App
89 7 0.8 438 6 US-10-282-798-1844 Sequence 1844, Ap
90 7 0.8 450 6 US-10-394-322A-15 Sequence 15, Appl
91 7 0.8 455 6 US-10-273-573-7309 Sequence 7309, Ap
92 7 0.8 536 6 US-10-461-791-10 Sequence 10, Appl
93 7 0.8 557 6 US-10-461-792-10 Sequence 439, App
94 7 0.8 575 6 US-10-326-956-439 Sequence 8, Appl
95 7 0.8 575 6 US-10-617-624-8 Sequence 8, Appl
96 7 0.8 575 6 US-10-620-061-8 Sequence 1266, Ap
97 7 0.8 586 6 US-10-408-765A-1266 Sequence 1266, Ap
98 7 0.8 596 6 US-10-273-573-5878 Sequence 5878, Ap
99 7 0.8 617 6 US-10-273-573-5880 Sequence 5880, Ap
100 7 0.8 617 6 US-10-273-573-7306 Sequence 7306, Ap
```

ALIGNMENTS

```
RESULT 1
US-10-357-820-28
; Sequence 28, Application US/10357820
; GENERAL INFORMATION:
; APPLICANT: Anderson, David W.;
; APPLICANT: Burgess, Catherine E.;
; APPLICANT: Casman, Stacie J.;
; APPLICANT: Gorman, Linda;
; APPLICANT: Ji, Weizhen;
; APPLICANT: Kekuda, Ramesh;
; APPLICANT: Li, Li;
; APPLICANT: Padigar, Muralidhara;
; APPLICANT: Patnirajan, Meera;
; APPLICANT: Pena, Carol E.A.;
; APPLICANT: Shenoy, Suresh G.;
; APPLICANT: Shimkete, Richard A.;
; APPLICANT: Stone, David J.;
; APPLICANT: Taupier, Raymond J.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-538B
; CURRENT APPLICATION NUMBER: US/10/357,820
; CURRENT FILING DATE: 2003-02-03
; PRIOR APPLICATION NUMBER: 09/679460
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: 09/730617
; PRIOR FILING DATE: 2000-12-05
; PRIOR APPLICATION NUMBER: 10/074978
; PRIOR FILING DATE: 2002-02-12
; PRIOR APPLICATION NUMBER: 10/138588
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: 60/387002
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/355099
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: 60/375579
; PRIOR FILING DATE: 2002-04-25
; PRIOR APPLICATION NUMBER: 60/393265
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 401825
; PRIOR FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: 60/381666
; PRIOR FILING DATE: 2002-05-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: CuraSeqList version 0.1
; SEQ ID NO 28
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-357-820-28
Alignment Scores: 7.55e-07 Length: 514
Pred. No.: 6
```

```
Score: 15.00 Matches: 15
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.68% Indels: 0
DB: 6 Gaps: 0
```

US-09-774-490-1 (1-2709) x US-10-357-820-28 (1-514)

OY 1028 AGAAGCTGTGTAATGAATGACAACTTCCTCAAGCTGCTG 1072

DB 274 ATGSetleuValAsnlyStrpThrPheleuylsAlaArgleu 288

RESULT 2

US-10-408-765A-2003

; Sequence 2003, Application US/10408765A

; GENERAL INFORMATION:

; APPLICANT: Ghosh, Soumitra S.

; APPLICANT: Fahy, Eoin D.

; APPLICANT: Zhang, Bing

; APPLICANT: Gibson, Bradford W.

; APPLICANT: Taylor, Steven W.

; APPLICANT: Glenn, Gary M.

; APPLICANT: Warnock, Dale E.

; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION

; FILE REFERENCE: 660088.465

; CURRENT APPLICATION NUMBER: US/10/408,765A

; CURRENT FILING DATE: 2003-04-04

; NUMBER OF SEQ ID NOS: 3077

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 2003

; LENGTH: 636

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-408-765A-2003

Alignment Scores:

Pred. No.: 7.43e-07 Length: 636

Score: 15.00 Matches: 15

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.68% Indels: 0

DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-408-765A-2003 (1-636)

OY 1028 AGAAGCTGTGTAATGAATGACAACTTCCTCAAGCTGCTG 1072

DB 164 ATGSetleuValAsnlyStrpThrPheleuylsAlaArgleu 178

RESULT 3

PCT-US03-09929-92

; Sequence 92, Application PCT/US0309929

; GENERAL INFORMATION:

; APPLICANT: Curagen Corporation, et al.

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD

; FILE REFERENCE: 21402-5738-061

; CURRENT APPLICATION NUMBER: PCT/US03/09929

; CURRENT FILING DATE: 2003-04-01

; PRIOR APPLICATION NUMBER: 60/368,996

; PRIOR FILING DATE: 2002-04-01

; PRIOR APPLICATION NUMBER: 60/369,980

; PRIOR FILING DATE: 2002-04-04

; PRIOR APPLICATION NUMBER: 60/370,381

; PRIOR FILING DATE: 2002-04-05

; PRIOR APPLICATION NUMBER: 60/370,969

; PRIOR FILING DATE: 2002-04-08

; PRIOR APPLICATION NUMBER: 60/371,002

; PRIOR FILING DATE: 2002-04-09

; PRIOR APPLICATION NUMBER: 60/372,002

; PRIOR FILING DATE: 2002-04-12

; PRIOR APPLICATION NUMBER: 60/384,297

; PRIOR FILING DATE: 2002-05-30

```

; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 92
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-09929-92

Alignment Scores:
Pred. No.: 7.57e-05 Length: 775
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US03-09929-92 (1-775)
OY 1778 TGTTCCTGCGCCGAGACCTTACTGTGCTGGATGCT 1816
DB 529 CysCysLeuAlaArgAspProTyrCysAlaIrrpAspGly 541

RESULT 4
PCT-US03-09929-94
; Sequence 94, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 94
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-09929-94

Alignment Scores:
Pred. No.: 7.57e-05 Length: 775
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
```

```

DB: 1 Gaps: 0
US-09-774-490-1 (1-2709) x PCT-US03-09929-94 (1-775)
OY 1778 TGTTCCTGCGCCGAGACCTTACTGTGCTGGATGCT 1816
DB 529 CysCysLeuAlaArgAspProTyrCysAlaIrrpAspGly 541

RESULT 5
PCT-US03-09929-96
; Sequence 96, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 96
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-09929-96

Alignment Scores:
Pred. No.: 7.57e-05 Length: 775
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US03-09929-96 (1-775)
OY 1778 TGTTCCTGCGCCGAGACCTTACTGTGCTGGATGCT 1816
DB 529 CysCysLeuAlaArgAspProTyrCysAlaIrrpAspGly 541

RESULT 6
PCT-US03-09929-98
; Sequence 98, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-573B-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
```

```

; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuroSeqList version 0.1
; SEQ ID NO 98
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-09929-98

Alignment Scores:
Pred. No.: 7,57e-05 Length: 775
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US03-09929-98 (1-775)

QY 1778 TGTTCCTGCGCCGAGACCTTACTGTGCTTGAGATGT 1816
DB 529 CysCysLeuAlaIArgAspProTyrCysAlaItrpAspGly 541

RESULT 7
; Sequence 97, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; FILE REFERENCE: 21402-5738-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shinkets, Richard
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
```

```

; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 97
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Domain search
US-10-369-072-97

Alignment Scores:
Pred. No.: 0.00819 Length: 431
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.23% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-97 (1-431)

QY 1040 AATAATGACAACTTCTCAAGCTGCTCG 1072
DB 221 AenlystipThrPhenleuysAlaIArgLeu 231

RESULT 8
; Sequence 4, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-5738-061
; CURRENT APPLICATION NUMBER: PCT/US03/09929
; CURRENT FILING DATE: 2003-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuroSeqList version 0.1
; SEQ ID NO 4
; LENGTH: 495
; TYPE: PRT
; ORGANISM: Homo sapiens
```

```
PCT-US03-09929-4
Alignment Scores:
Pred. No.: 0.0825 Length: 495
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US03-09929-4 (1-495)
QY 917 GATGACAAAGTACTTTTCTTCGCTGAA 946
Db 232 AspaaplyValtyrphhepheargGlu 241

RESULT 9
PCT-US03-09929-2
; Sequence 2, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-5738-061
; CURRENT FILING DATE: 2003-04-01
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: 60/402,207
; PRIOR FILING DATE: 2002-08-09
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: CuraseqList version 0.1
; SEQ ID NO 2
; LENGTH: 634
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US03-09929-2

Alignment Scores:
Pred. No.: 0.0809 Length: 634
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US03-09929-2 (1-634)
QY 917 GATGACAAAGTACTTTTCTTCGCTGAA 946
Db 212 AspaaplyValtyrphhepheargGlu 221

RESULT 10
PCT-US03-09929-8
; Sequence 8, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.

TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
FILE REFERENCE: 21402-5738-061
CURRENT FILING DATE: 2003-04-01
PRIOR FILING DATE: 2002-04-01
PRIOR APPLICATION NUMBER: 60/368,996
PRIOR FILING DATE: 2002-04-01
PRIOR APPLICATION NUMBER: 60/369,980
PRIOR FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: 60/370,381
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: 60/370,969
PRIOR FILING DATE: 2002-04-08
PRIOR APPLICATION NUMBER: 60/371,002
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/372,002
PRIOR FILING DATE: 2002-04-12
PRIOR APPLICATION NUMBER: 60/384,297
PRIOR FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: 60/386,816
PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: 60/389,123

TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
FILE REFERENCE: 21402-5738-061
CURRENT FILING DATE: 2003-04-01
PRIOR FILING DATE: 2002-04-01
PRIOR APPLICATION NUMBER: 60/368,996
PRIOR FILING DATE: 2002-04-01
PRIOR APPLICATION NUMBER: 60/369,980
PRIOR FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: 60/370,381
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: 60/370,969
PRIOR FILING DATE: 2002-04-08
PRIOR APPLICATION NUMBER: 60/371,002
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/372,002
PRIOR FILING DATE: 2002-04-12
PRIOR APPLICATION NUMBER: 60/384,297
PRIOR FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: 60/386,816
PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: 60/389,123

US-09-774-490-1 (1-2709) x PCT-US03-09929-8 (1-638)
QY 917 GATGACAAAGTACTTTTCTTCGCTGAA 946
Db 214 AspaaplyValtyrphhepheargGlu 223

RESULT 11
PCT-US03-09929-10
; Sequence 10, Application PC/TUS0309929
; GENERAL INFORMATION:
; APPLICANT: Curagen Corporation, et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-5738-061
; CURRENT FILING DATE: 2003-04-01
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/368,996
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 60/369,980
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 60/370,381
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: 60/370,969
; PRIOR FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: 60/371,002
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/372,002
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/384,297
; PRIOR FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: 60/386,816
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/389,123
```

;; PRIOR FILING DATE: 2002-06-13
;; PRIOR APPLICATION NUMBER: 60/402,207
;; PRIOR FILING DATE: 2002-08-09
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 179
;; SOFTWARE: CuraseqList version 0.1

;; SEQ ID NO 10
;; LENGTH: 638
;; TYPE: PRT
;; ORGANISM: Homo sapiens
PCT-US03-09929-10

Alignment Scores:
Pred. No.: 0.0809
Score: 10.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.12%
DB: 1

US-09-774-490-1 (1-2709) x PCT-US03-09929-10 (1-638)
QY 917 GATGACAAAGTATCTTTCTTCCTCCGTGA 946
Db 214 AspAspIyValIyTrPhepPheArGlu 223

RESULT 12
PCT-US03-09929-6
Sequence 6, Application PC/TUS0309929
GENERAL INFORMATION:
APPLICANT: Curagen Corporation, et al.
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
FILE REFERENCE: 21402-573B-061
CURRENT APPLICATION NUMBER: PCT/US03/09929
CURRENT FILING DATE: 2003-04-01

PRIOR APPLICATION NUMBER: 60/368,996
PRIOR FILING DATE: 2002-04-01
PRIOR APPLICATION NUMBER: 60/369,980
PRIOR FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: 60/370,381
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: 60/370,969
PRIOR FILING DATE: 2002-04-08
PRIOR APPLICATION NUMBER: 60/371,002
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/372,002
PRIOR FILING DATE: 2002-04-12
PRIOR APPLICATION NUMBER: 60/384,297
PRIOR FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: 60/386,816
PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: 60/389,123
PRIOR FILING DATE: 2002-06-13
PRIOR APPLICATION NUMBER: 60/402,207
PRIOR FILING DATE: 2002-08-09
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 179
SOFTWARE: CuraseqList version 0.1
SEQ ID NO 6
LENGTH: 833
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US03-09929-6

Alignment Scores:

Pred. No.: 0.0793
Score: 10.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.12%
DB: 1

US-09-774-490-1 (1-2709) x PCT-US03-09929-6 (1-833)

QY 917 GATGACAAAGTATCTTTCTTCCTCCGTGA 946
Db 232 AspAspIyValIyTrPhepPheArGlu 241

RESULT 13
PCT-US03-09929-12
Sequence 12, Application PC/TUS0309929
GENERAL INFORMATION:
APPLICANT: Curagen Corporation, et al.
TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
FILE REFERENCE: 21402-573B-061
CURRENT APPLICATION NUMBER: PCT/US03/09929
CURRENT FILING DATE: 2003-04-01

PRIOR APPLICATION NUMBER: 60/368,996
PRIOR FILING DATE: 2002-04-01
PRIOR APPLICATION NUMBER: 60/369,980
PRIOR FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: 60/370,381
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: 60/370,969
PRIOR FILING DATE: 2002-04-08
PRIOR APPLICATION NUMBER: 60/371,002
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/372,002
PRIOR FILING DATE: 2002-04-12
PRIOR APPLICATION NUMBER: 60/384,297
PRIOR FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: 60/386,816
PRIOR FILING DATE: 2002-06-07
PRIOR APPLICATION NUMBER: 60/389,123
PRIOR FILING DATE: 2002-06-13
PRIOR APPLICATION NUMBER: 60/402,207
PRIOR FILING DATE: 2002-08-09
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 179
SOFTWARE: CuraseqList version 0.1
SEQ ID NO 12
LENGTH: 833
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US03-09929-12

Alignment Scores:
Pred. No.: 0.0793
Score: 10.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.12%
DB: 1

US-09-774-490-1 (1-2709) x PCT-US03-09929-12 (1-833)

QY 1043 AATGACAACTTCTCTCAAGCTGCTG 1072
Db 274 LysTrpTrpTrpPheLeuLysAlaArgLeu 283

RESULT 14
US-10-369-072-14
Sequence 14, Application US/10369072
GENERAL INFORMATION:
APPLICANT: Alsobrook II, John P
APPLICANT: Spaderna, Stephen K
APPLICANT: Tchernev, Velizar
APPLICANT: Liu, Xiaohong
APPLICANT: Shenoy, Suresh
APPLICANT: Spytek, Kimberly
APPLICANT: Zernusen, Bryan
APPLICANT: Patutajan, Meera
APPLICANT: Taupier, Raymond T
APPLICANT: Rastelli, Luca
APPLICANT: Grosse, William M
APPLICANT: Szerkes, Edward S


```
/ APPLICANT: Lepley, Denise M
/ APPLICANT: Shen, Lei
/ APPLICANT: Burgess, Catherine E
/ APPLICANT: Shimkets, Richard
/ APPLICANT: Padigaru, Muralidhara
/ TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-050 CON2
/ CURRENT APPLICATION NUMBER: US/10/369,072
/ CURRENT FILING DATE: 2003-02-18
/ PRIOR APPLICATION NUMBER: 10/174,372
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 09/898,994
/ PRIOR FILING DATE: 2001-07-03
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,586
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,722
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/218,622
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 60/218,992
/ PRIOR FILING DATE: 2000-07-17
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 14
/ LENGTH: 833
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-10-369-072-14

Alignment Scores:
Pred. No.: 0.0793 Length: 833
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-14 (1-833)

QY 1043 AATGACACATCTCTCAAGCTGCTG 1072
Db 274 LysTPTnThrPheLeuIysAlaArgLeu 283

RESULT 15
/ Sequence 16, Application US/10369072
/ GENERAL INFORMATION:
/ APPLICANT: Alsobrook II, John P
/ APPLICANT: Spaderna, Stephen K
/ APPLICANT: Tchenev, Vellizar
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Shenoy, Suresh
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zetnusen, Bryan
/ APPLICANT: Patruzjan, Meera
/ APPLICANT: Taupier, Raymond T
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Grose, William M
/ APPLICANT: Szerkes, Edward S
/ APPLICANT: Lepley, Denise M
/ APPLICANT: Shen, Lei
/ APPLICANT: Burgess, Catherine E
/ APPLICANT: Shimkets, Richard
/ APPLICANT: Padigaru, Muralidhara
```

```
/ TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-050 CON2
/ CURRENT APPLICATION NUMBER: US/10/369,072
/ CURRENT FILING DATE: 2003-02-18
/ PRIOR APPLICATION NUMBER: 10/174,372
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 09/898,994
/ PRIOR FILING DATE: 2001-07-03
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,586
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,722
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/218,622
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 60/218,992
/ PRIOR FILING DATE: 2000-07-17
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 16
/ LENGTH: 833
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ US-10-369-072-16

Alignment Scores:
Pred. No.: 0.0793 Length: 833
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-16 (1-833)

QY 917 GATGACAAATGATCTTTCTCGCGAA 946
Db 232 AspaAsplysValIyTyrPhePheArgGlu 241

RESULT 16
/ Sequence 18, Application US/10369072
/ GENERAL INFORMATION:
/ APPLICANT: Alsobrook II, John P
/ APPLICANT: Spaderna, Stephen K
/ APPLICANT: Tchenev, Vellizar
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Shenoy, Suresh
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zetnusen, Bryan
/ APPLICANT: Patruzjan, Meera
/ APPLICANT: Taupier, Raymond T
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Grose, William M
/ APPLICANT: Szerkes, Edward S
/ APPLICANT: Lepley, Denise M
/ APPLICANT: Shen, Lei
/ APPLICANT: Burgess, Catherine E
/ APPLICANT: Shimkets, Richard
/ APPLICANT: Padigaru, Muralidhara
/ TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-050 CON2
/ CURRENT APPLICATION NUMBER: US/10/369,072
/ CURRENT FILING DATE: 2003-02-18
/ PRIOR APPLICATION NUMBER: 10/174,372
```

```
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 09/898,994
/ PRIOR FILING DATE: 2001-07-03
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,586
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,722
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/218,622
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 60/218,992
/ PRIOR FILING DATE: 2000-07-17
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: Patentln Ver. 2.1
/ SEQ ID NO 18
/ LENGTH: 833
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-369-072-18

Alignment Scores:
Pred. No.: 0.0793 Length: 833
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-18 (1-833)

QY 917 GATGACAAAGTATCTTTCTTCCTCGTGA 946
Db 232 AspAspIyValTyIrhPhepneArGlu 241

RESULT 17
US-10-369-072-47
/ Sequence 47, Application US/10369072
/ GENERAL INFORMATION:
/ APPLICANT: Alsobrook II, John P
/ APPLICANT: Spaderma, Stephen K
/ APPLICANT: Tchernev, Velizar
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Shenoy, Suresh
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zehnusen, Bryan
/ APPLICANT: Patlurajan, Meera
/ APPLICANT: Taupier, Raymond T
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Grosse, William M
/ APPLICANT: Szerkeres, Edward S
/ APPLICANT: Lepley, Denise M
/ APPLICANT: Shen, Lei
/ APPLICANT: Burgess, Catherine E
/ APPLICANT: Shinkets, Richard
/ APPLICANT: Padigaru, Muralidhara
/ TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-050 CON2
/ CURRENT APPLICATION NUMBER: US/10/369,072
/ CURRENT FILING DATE: 2003-02-18
/ PRIOR APPLICATION NUMBER: 10/174,372
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 09/898,994
/ PRIOR FILING DATE: 2001-07-03
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2000-07-03
```

```
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,586
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,722
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/218,622
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 60/218,992
/ PRIOR FILING DATE: 2000-07-17
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: Patentln Ver. 2.1
/ SEQ ID NO 47
/ LENGTH: 834
/ TYPE: PRT
/ ORGANISM: Mus musculus
US-10-369-072-47

Alignment Scores:
Pred. No.: 0.0793 Length: 834
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-47 (1-834)

QY 1043 AAATGACAACTTCTCAAGCTGCTCG 1072
Db 274 LysTpThrThrhPheulysAlaArgLeu 283

RESULT 18
US-10-369-072-46
/ Sequence 46, Application US/10369072
/ GENERAL INFORMATION:
/ APPLICANT: Alsobrook II, John P
/ APPLICANT: Spaderma, Stephen K
/ APPLICANT: Tchernev, Velizar
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Shenoy, Suresh
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zehnusen, Bryan
/ APPLICANT: Patlurajan, Meera
/ APPLICANT: Taupier, Raymond T
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Grosse, William M
/ APPLICANT: Szerkeres, Edward S
/ APPLICANT: Lepley, Denise M
/ APPLICANT: Shen, Lei
/ APPLICANT: Burgess, Catherine E
/ APPLICANT: Shinkets, Richard
/ APPLICANT: Padigaru, Muralidhara
/ TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-050 CON2
/ CURRENT APPLICATION NUMBER: US/10/369,072
/ CURRENT FILING DATE: 2003-02-18
/ PRIOR APPLICATION NUMBER: 10/174,372
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 09/898,994
/ PRIOR FILING DATE: 2001-07-03
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
```

```

; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 963
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-072-46

Alignment Scores:
Pred. No.: 0.0784 Length: 963
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-46 (1-963)

QY 917 GATGACAAAGTACTTTCTTCCTCGTGA 946
DB 362 AepAepLyValTyPhephepheaTgGlu 371

RESULT 19
US-10-376-397A-42
; Sequence 42, Application US/10376397A
; GENERAL INFORMATION:
; APPLICANT: Ana L.T.O. NACIMENTO et al.
; TITLE OF INVENTION: SURFACE PROTEINS OF LEPTOSPIRA
; FILE REFERENCE: FAPESP 204.1 US (10303399)
; CURRENT APPLICATION NUMBER: US/10/376,397A
; PRIOR FILING DATE: 2003-02-28
; NUMBER OF SEQ ID NOS: 46
; SEQ ID NO 42
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Leptospira interrogans serovar copenhageni
US-10-376-397A-42

Alignment Scores:
Pred. No.: 0.861 Length: 350
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-376-397A-42 (1-350)

QY 798 CTATTCGGAACCTTTGGGACACACC 824
DB 188 LeuSerSerGluLeuLeuGlyThrThr 196

RESULT 20
US-10-369-072-48
; Sequence 48, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Zernusen, Bryan
```

```

; APPLICANT: Patutajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shinkets, Richard
; APPLICANT: Padigaru, Murilidhara
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; PRIOR FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48
; LENGTH: 510
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-072-48

Alignment Scores:
Pred. No.: 0.837 Length: 510
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-48 (1-510)

QY 1784 CTCGCCGAGACCTTACTGTGCTTG 1810
DB 188 LeuAlaArgAspProTyCysAlaTrp 196

RESULT 21
US-10-369-072-39
; Sequence 39, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Patutajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkeres, Edward S
```

```

; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigar, Muralidhara
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; PRIORITY FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 39
; LENGTH: 782
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-369-072-39

Alignment Scores:
Pred. No.: 0 81      Length: 782
Score: 9.00      Matches: 9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 1.01%      Indels: 0
DB: 6      Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-39 (1-782)

QY 1784 CTCGCCGAGACCTTACTGTGCTTG 1810
DB 483 LeuAlaArgAspProtyrCysAlaIrrp 491

RESULT 22
US-10-369-072-38
; Sequence 38, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsbrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Vellizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Splytek, Kimberly
; APPLICANT: Zehnsen, Bryan
; APPLICANT: Patuturajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grose, William M
; APPLICANT: Szerkes, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shimkets, Richard
; APPLICANT: Padigar, Muralidhara
```

```

; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; PRIORITY FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 893
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-072-38

Alignment Scores:
Pred. No.: 0 802      Length: 893
Score: 9.00      Matches: 9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 1.01%      Indels: 0
DB: 6      Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-38 (1-893)

QY 1784 CTCGCCGAGACCTTACTGTGCTTG 1810
DB 593 LeuAlaArgAspProtyrCysAlaIrrp 601

RESULT 23
US-10-326-956-1715
; Sequence 1715, Application US/10326956
; GENERAL INFORMATION:
; APPLICANT: Bauer et al.
; TITLE OF INVENTION: Protein Complexes and Methods for their Use
; FILE REFERENCE: 220615
; CURRENT APPLICATION NUMBER: US/10/326,956
; PRIORITY FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: EP 01 130 253.6
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 3282
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1715
; LENGTH: 1244
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-326-956-1715

Alignment Scores:
Pred. No.: 0 783      Length: 1244
Score: 9.00      Matches: 9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 1.01%      Indels: 0
DB: 6      Gaps: 0
```

US-09-774-490-1 (1-2709) x US-10-326-956-1715 (1-1244)

QY 380 CTTTGATGAGGACGAGTAGCTG 406
|||||
Db 607 LeuLeuAspGluGluArgSerArgLeu 615

RESULT 24

US-10-408-765A-2436
; Sequence 2436, Application US/10408765A
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Fahy, Eoin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; FILE REFERENCE: 660088.465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2436
; LENGTH: 3647
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-408-765A-2436

Alignment Scores:

Pred. No.: 0.721 Length: 3647
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-408-765A-2436 (1-3647)

QY 67 CGTGAGTCGACACCCCTGACGTGACA 41
|||||
Db 2233 ArgGlySerThrProLeuAspValAla 2241

RESULT 25

US-60-487-610-1645
; Sequence 1645, Application US/60487610
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele
; APPLICANT: HUANG, Hongjin
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: LIVER FIBROSIS IN HEPATITIS C VIRUS-INFECTED SUBJECTS,
; FILE REFERENCE: CL001469
; CURRENT APPLICATION NUMBER: US/60/487,610
; CURRENT FILING DATE: 2003-07-17
; NUMBER OF SEQ ID NOS: 97101
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1645
; LENGTH: 275
; TYPE: PRF
; ORGANISM: Homo sapiens
US-60-487-610-1645

Alignment Scores:

Pred. No.: 8.92 Length: 275
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 7 Gaps: 0

US-09-774-490-1 (1-2709) x US-60-487-610-1645 (1-275)

QY 1780 TTGCTCGCGCGAGACCTTACTG 1803
|||||
Db 22 LeuProArgProArgProLeuLeu 29

RESULT 26

US-10-292-798-748
; Sequence 748, Application US/10292798
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 08435/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 748
; LENGTH: 312
; TYPE: PRF
; ORGANISM: Homo sapiens
US-10-292-798-748

Alignment Scores:

Pred. No.: 8.84 Length: 312
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-292-798-748 (1-312)

QY 1770 GTGCTGAGTGTGCTCGCCGAG 1793
|||||
Db 225 ValLeuSerValAlaSerProGlu 232

RESULT 27

US-10-369-072-10
; Sequence 10, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tchernev, Velizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernusen, Bryan
; APPLICANT: Patnrajan, Meera
; APPLICANT: Taupier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkeres, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shinkels, Richard
; APPLICANT: Padigaru, Muralidhara
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT APPLICATION NUMBER: US/10/369,072
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03

```
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,586
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,722
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/218,622
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 60/218,992
/ PRIOR FILING DATE: 2000-07-17
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 10
/ LENGTH: 754
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-369-072-10

Alignment Scores:
Pred. No.: 8 27      length: 754
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.89%      Indels: 0
DB: 6      Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-10 (1-754)

QY 1043 AATGACACAACTTCCTCAAGCT 1066
DB 275 LyeTTPThrPhenLeuYsAla 282

RESULT 28
US-10-369-072-8
/ Sequence 8, Application US/10369072
/ GENERAL INFORMATION:
/ APPLICANT: Alsobrook II, John P
/ APPLICANT: Spaderna, Stephen K
/ APPLICANT: Tchernev, Velizar
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Shenoy, Suresh
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zehnusen, Bryan
/ APPLICANT: Paturajan, Meera
/ APPLICANT: Taupier, Raymond T
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Grose, William M
/ APPLICANT: Szerkres, Edward S
/ APPLICANT: Lepley, Denise M
/ APPLICANT: Shen, Lei
/ APPLICANT: Burgess, Catherine E
/ APPLICANT: Shinkets, Richard
/ APPLICANT: Padigaru, Muralidhara
/ TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-050 CON2
/ CURRENT APPLICATION NUMBER: US/10/369,072
/ PRIOR FILING DATE: 2003-02-18
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2001-07-03
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
```

```
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,586
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,722
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/218,622
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 60/218,992
/ PRIOR FILING DATE: 2000-07-17
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 8
/ LENGTH: 756
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-369-072-8

Alignment Scores:
Pred. No.: 8 27      length: 756
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.89%      Indels: 0
DB: 6      Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-8 (1-756)

QY 1043 AATGACACAACTTCCTCAAGCT 1066
DB 282 LyeTTPThrPhenLeuYsAla 289

RESULT 29
US-10-369-072-36
/ Sequence 36, Application US/10369072
/ GENERAL INFORMATION:
/ APPLICANT: Alsobrook II, John P
/ APPLICANT: Spaderna, Stephen K
/ APPLICANT: Tchernev, Velizar
/ APPLICANT: Liu, Xiaohong
/ APPLICANT: Shenoy, Suresh
/ APPLICANT: Spytek, Kimberly
/ APPLICANT: Zehnusen, Bryan
/ APPLICANT: Paturajan, Meera
/ APPLICANT: Taupier, Raymond T
/ APPLICANT: Rastelli, Luca
/ APPLICANT: Grose, William M
/ APPLICANT: Szerkres, Edward S
/ APPLICANT: Lepley, Denise M
/ APPLICANT: Shen, Lei
/ APPLICANT: Burgess, Catherine E
/ APPLICANT: Shinkets, Richard
/ APPLICANT: Padigaru, Muralidhara
/ TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
/ FILE REFERENCE: 21402-050 CON2
/ CURRENT APPLICATION NUMBER: US/10/369,072
/ PRIOR FILING DATE: 2003-02-18
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2002-06-17
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2001-07-03
/ PRIOR APPLICATION NUMBER: 60/215,854
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,856
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/215,902
/ PRIOR FILING DATE: 2000-07-03
/ PRIOR APPLICATION NUMBER: 60/216,585
/ PRIOR FILING DATE: 2000-07-07
/ PRIOR APPLICATION NUMBER: 60/216,586
/ PRIOR FILING DATE: 2000-07-07
```

PRIOR APPLICATION NUMBER: 60/216,722
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/218,622
PRIOR FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 60/218,992
PRIOR FILING DATE: 2000-07-17
Remaining prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 100
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 36
LENGTH: 760
TYPE: PRT
ORGANISM: Mus musculus
US-10-369-072-36

Alignment Scores:
Pred. No.: 8.26 Length: 760
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-36 (1-760)

Qy 578 ACTCACTTGACGCTGTGAACG 601
Db 137 ThrHSeuTyRAlAcyselYthr 144

RESULT 30
US-10-293-244-3947
Sequence 3947, Application US/10293244
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
FILE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-029
CURRENT APPLICATION NUMBER: US/10/293,244
CURRENT FILING DATE: 2002-11-12
PRIOR APPLICATION NUMBER: Not Yet Assigned
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 09/728,422
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 09/593,325
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/663,561
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/654,936
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 09/520,325
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/598,075
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 3960
SOFTWARE: Custom
SEQ ID NO 3947
LENGTH: 761
TYPE: PRT
ORGANISM: Homo sapiens
US-10-293-244-3947

Alignment Scores:
Pred. No.: 8.26 Length: 761
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-293-244-3947 (1-761)
Qy 1043 AATGACACATTCCTCAAGCT 1066
Db 282 LysTrpThrPheLeuYsAla 289

RESULT 31
US-10-369-072-35
Sequence 35, Application US/10369072
GENERAL INFORMATION:
APPLICANT: Alsobrook II, John P
APPLICANT: Spaderna, Stephen K
APPLICANT: Tchenev, Velizar
APPLICANT: Liu, Xiaohong
APPLICANT: Shenoy, Suresh
APPLICANT: Spytek, Kimberly
APPLICANT: Zernusen, Bryan
APPLICANT: Paturajan, Meera
APPLICANT: Taupier, Raymond T
APPLICANT: Rastelli, Luca
APPLICANT: Grosse, William M
APPLICANT: Szerkeres, Edward S
APPLICANT: Lepley, Denise M
APPLICANT: Shen, Lei
APPLICANT: Burgess, Catherine E
APPLICANT: Shinkels, Richard
APPLICANT: Padigaru, Muralidhara
TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-050 CON2
CURRENT APPLICATION NUMBER: US/10/369,072
CURRENT FILING DATE: 2003-02-18
PRIOR APPLICATION NUMBER: 10/174,372
PRIOR FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: 09/898,994
PRIOR FILING DATE: 2001-07-03
PRIOR APPLICATION NUMBER: 60/215,854
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/215,856
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/215,902
PRIOR FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/216,585
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/216,586
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/216,722
PRIOR FILING DATE: 2000-07-07
PRIOR APPLICATION NUMBER: 60/218,622
PRIOR FILING DATE: 2000-07-17
PRIOR APPLICATION NUMBER: 60/218,992
PRIOR FILING DATE: 2000-07-17
Remaining prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 100
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 35
LENGTH: 762
TYPE: PRT
ORGANISM: Homo sapiens
US-10-369-072-35

Alignment Scores:
Pred. No.: 8.26 Length: 762
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-35 (1-762)

Qy 1043 AATGACACATTCCTCAAGCT 1066
Db 282 LysTrpThrPheLeuYsAla 289

```

RESULT 32
US-10-369-072-49
; Sequence 49, Application US/10369072
; GENERAL INFORMATION:
; APPLICANT: Alsobrook II, John P
; APPLICANT: Spaderna, Stephen K
; APPLICANT: Tochernev, Vellizar
; APPLICANT: Liu, Xiaohong
; APPLICANT: Shenoy, Suresh
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zernhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Tauplier, Raymond T
; APPLICANT: Rastelli, Luca
; APPLICANT: Grosse, William M
; APPLICANT: Szerkerez, Edward S
; APPLICANT: Lepley, Denise M
; APPLICANT: Shen, Lei
; APPLICANT: Burgess, Catherine E
; APPLICANT: Shikets, Richard
; APPLICANT: Padigar, Muralidhara
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-050 CON2
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: 10/174,372
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 09/898,994
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/215,854
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,856
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/215,902
; PRIOR FILING DATE: 2000-07-03
; PRIOR APPLICATION NUMBER: 60/216,585
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,586
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/216,722
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,622
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 60/218,992
; PRIOR FILING DATE: 2000-07-17
; Remaining Prior Application data removed - See File Wrapper or PALM
; NUMBER OF SEQ ID NOS: 100
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 49
; LENGTH: 838
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-369-072-49

Alignment Scores:
Score: 8.2 Length: 838
Pred. No.: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-49 (1-838)

OY 578 ACTCATTTGTACGCGCTGTGGAGC 601
|||||
Db 128 TTTTATSTeutyAlAcysgIlyThr 135

RESULT 33
US-10-273-573-9338
; Sequence 9338, Application US/10273573
; GENERAL INFORMATION:

```

```

APPLICANT: Hysq, Inc
TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-065
CURRENT APPLICATION NUMBER: US/10/273,573
CURRENT FILING DATE: 2002-10-18
PRIOR APPLICATION NUMBER: 09/522,929
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: 09/770,160
PRIOR FILING DATE: 2001-01-26
NUMBER OF SEQ ID NOS: 10994
SOFTWARE: Custom
SEQ ID NO 9338
LENGTH: 1074
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (598)..(626)
OTHER INFORMATION: PRECURSOR GLYCOPROTEIN SIGNAL RE domain identified by
OTHER INFORMATION: SMARTX, accession number PD01719A, p-value=6.100e-15, raw score=
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (58)..(468)
OTHER INFORMATION: Sema domain identified by Pfam, accession name Sema, E-value=
OTHER INFORMATION: 4.6e-196, Pfam score of 664.7
US-10-273-9338

Alignment Scores:
Pred. No.:      8.05      Length:      1074
Score:          8.00      Matches:      8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:      0.89%      Indels:      0
DB:              Gaps:      0

US-09-774-490-1 (1-2709) x US-10-273-9338 (1-1074)

QY      929 TACTTTTCTTCGCGAAATGCA 952
      |||||||
Db      230 TyrPhePhePheArgGluAsnAla 237

RESULT 34
US-10-369-072-41
Sequence 41, Application US/10369072
GENERAL INFORMATION:
APPLICANT: Alsobrook II, John P
APPLICANT: Spaderna, Stephen K
APPLICANT: Tchenev, Velizar
APPLICANT: Liu, Xiaohong
APPLICANT: Shenoy, Suresh
APPLICANT: Spytek, Kimberly
APPLICANT: Zerhusen, Bryan
APPLICANT: Patturajan, Meera
APPLICANT: Taupier, Raymond T
APPLICANT: Rastelli, Luca
APPLICANT: Grosse, William M
APPLICANT: Szerkess, Edward S
APPLICANT: Lepley, Denise M
APPLICANT: Shen, Lei
APPLICANT: Burgess, Catherine E
APPLICANT: Shimkels, Richard
APPLICANT: Padiguru, Muralidhara
TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
FILE REFERENCE: 21402-050 CON2
CURRENT APPLICATION NUMBER: US/10/369,072
CURRENT FILING DATE: 2003-02-18
PRIOR APPLICATION NUMBER: 10/174,372
PRIOR FILING DATE: 2002-06-17
PRIOR APPLICATION NUMBER: 09/898,994
PRIOR FILING DATE: 2001-07-03
PRIOR APPLICATION NUMBER: 60/215,854
PRIOR FILING DATE: 2000-07-03

```


;; PRIOR APPLICATION NUMBER: 60/215,856
;; PRIOR FILING DATE: 2000-07-03
;; PRIOR APPLICATION NUMBER: 60/215,902
;; PRIOR FILING DATE: 2000-07-03
;; PRIOR APPLICATION NUMBER: 60/216,585
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/216,586
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/216,722
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/218,622
;; PRIOR FILING DATE: 2000-07-17
;; PRIOR APPLICATION NUMBER: 60/218,992
;; PRIOR FILING DATE: 2000-07-17
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 100
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 41
;; LENGTH: 1074
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-10-369-072-41

Alignment Scores:
Pred. No.: 8.05 Length: 1074
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-41 (1-1074)

QY 929 TACTTTTCTCGGAAATGCA 952
|||
Db 230 TYRPNPNEPNEARGLUASMA 237

RESULT 35

;; Sequence 44, Application US/10369072
;; GENERAL INFORMATION:
;; APPLICANT: Alsobrook II, John P
;; APPLICANT: Spaderna, Stephen K
;; APPLICANT: Tchernev, Velizar
;; APPLICANT: Liu, Xiaohong
;; APPLICANT: Shenoy, Suresh
;; APPLICANT: Zernusen, Bryan
;; APPLICANT: Patnrajan, Meera
;; APPLICANT: Taupier, Raymond T
;; APPLICANT: Rastelli, Luca
;; APPLICANT: Grosse, William M
;; APPLICANT: Szerkerez, Edward S
;; APPLICANT: Lepley, Denise M
;; APPLICANT: Shen, Lei
;; APPLICANT: Burgess, Catherine E
;; APPLICANT: Shinkets, Richard
;; APPLICANT: Padigaru, Muralidhara
;; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
;; FILE REFERENCE: 21402-050 CON2
;; CURRENT APPLICATION NUMBER: US/10/369,072
;; PRIOR FILING DATE: 2003-02-18
;; PRIOR APPLICATION NUMBER: 10/174,372
;; PRIOR FILING DATE: 2002-06-17
;; PRIOR APPLICATION NUMBER: 09/898,994
;; PRIOR FILING DATE: 2001-07-03
;; PRIOR APPLICATION NUMBER: 60/215,854
;; PRIOR FILING DATE: 2000-07-03
;; PRIOR APPLICATION NUMBER: 60/215,856
;; PRIOR FILING DATE: 2000-07-03
;; PRIOR APPLICATION NUMBER: 60/215,902
;; PRIOR FILING DATE: 2000-07-03
;; PRIOR APPLICATION NUMBER: 60/216,585

;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/216,586
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/216,722
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/218,622
;; PRIOR FILING DATE: 2000-07-17
;; PRIOR APPLICATION NUMBER: 60/218,992
;; PRIOR FILING DATE: 2000-07-17
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 100
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 44
;; LENGTH: 1074
;; TYPE: PRF
;; ORGANISM: Homo sapiens
US-10-369-072-44

Alignment Scores:
Pred. No.: 8.05 Length: 1074
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-44 (1-1074)

QY 929 TACTTTTCTCGGAAATGCA 952
|||
Db 230 TYRPNPNEPNEARGLUASMA 237

RESULT 36

;; Sequence 43, Application US/10369072
;; GENERAL INFORMATION:
;; APPLICANT: Alsobrook II, John P
;; APPLICANT: Spaderna, Stephen K
;; APPLICANT: Tchernev, Velizar
;; APPLICANT: Liu, Xiaohong
;; APPLICANT: Shenoy, Suresh
;; APPLICANT: Zernusen, Bryan
;; APPLICANT: Patnrajan, Meera
;; APPLICANT: Taupier, Raymond T
;; APPLICANT: Rastelli, Luca
;; APPLICANT: Grosse, William M
;; APPLICANT: Szerkerez, Edward S
;; APPLICANT: Lepley, Denise M
;; APPLICANT: Shen, Lei
;; APPLICANT: Burgess, Catherine E
;; APPLICANT: Shinkets, Richard
;; APPLICANT: Padigaru, Muralidhara
;; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
;; FILE REFERENCE: 21402-050 CON2
;; CURRENT APPLICATION NUMBER: US/10/369,072
;; PRIOR FILING DATE: 2003-02-18
;; PRIOR APPLICATION NUMBER: 10/174,372
;; PRIOR FILING DATE: 2002-06-17
;; PRIOR APPLICATION NUMBER: 09/898,994
;; PRIOR FILING DATE: 2001-07-03
;; PRIOR APPLICATION NUMBER: 60/215,854
;; PRIOR FILING DATE: 2000-07-03
;; PRIOR APPLICATION NUMBER: 60/215,856
;; PRIOR FILING DATE: 2000-07-03
;; PRIOR APPLICATION NUMBER: 60/215,902
;; PRIOR FILING DATE: 2000-07-03
;; PRIOR APPLICATION NUMBER: 60/216,585
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/216,586
;; PRIOR FILING DATE: 2000-07-07
;; PRIOR APPLICATION NUMBER: 60/216,722
;; PRIOR FILING DATE: 2000-07-07

```
/ PRIOR APPLICATION NUMBER: 60/218,622
/ PRIOR FILING DATE: 2000-07-17
/ PRIOR APPLICATION NUMBER: 60/218,992
/ PRIOR FILING DATE: 2000-07-17
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 100
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 43
/ LENGTH: 1077
/ TYPE: PRF
/ ORGANISM: Mus musculus
US-10-369-072-43

Alignment Scores:
Pred. No.: 8.05 Length: 1077
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-369-072-43 (1-1077)
QY 929 TACTTTTCTCCGTGAATAATGCA 952
DB 230 TyPhePhePheArgGluSerAla 237

RESULT 37
US-10-408-765A-26
/ Sequence 26, Application US/10408765A
/ GENERAL INFORMATION:
/ APPLICANT: Ghosh, Soumitra S.
/ APPLICANT: Fahy, Bojn D.
/ APPLICANT: Zhang, Bing
/ APPLICANT: Gibson, Bradford W.
/ APPLICANT: Taylor, Steven W.
/ APPLICANT: Glenn, Gary M.
/ APPLICANT: Warnock, Dale E.
/ TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
/ TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
/ FILE REFERENCE: 660088,465
/ CURRENT APPLICATION NUMBER: US/10/408,765A
/ CURRENT FILING DATE: 2003-04-04
/ NUMBER OF SEQ ID NOS: 3077
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 26
/ LENGTH: 5032
/ TYPE: PRF
/ ORGANISM: Homo sapiens
US-10-408-765A-26

Alignment Scores:
Pred. No.: 7.16 Length: 5032
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-408-765A-26 (1-5032)
QY 64 GGGTCGACACCCCTCGACGTGCA 41
DB 2268 GlySerThrProLeuAspValAla 2275

RESULT 38
US-10-462-850-2385
/ Sequence 2385, Application US/10462850
/ GENERAL INFORMATION:
/ APPLICANT: Proteom Ltd
/ TITLE OF INVENTION: Complementary peptide ligands from the human genome
/ FILE REFERENCE: Human patent
/ CURRENT APPLICATION NUMBER: US/10/462,850
/ CURRENT FILING DATE: 2003-06-17
/ NUMBER OF SEQ ID NOS: 4203
/ SOFTWARE: ProPatent version 1.0
/ SEQ ID NO 2399
/ LENGTH: 10

CURRENT FILING DATE: 2003-06-17
NUMBER OF SEQ ID NOS: 4203
SOFTWARE: ProPatent version 1.0
SEQ ID NO 2385
LENGTH: 10
TYPE: PRF
ORGANISM: Homo Sapiens
FEATURE:
OTHER INFORMATION: sequence located in C8A at 82-91 and may interact with Sequence 2
US-10-462-850-2385

Alignment Scores:
Pred. No.: 117 Length: 10
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-462-850-2385 (1-10)
QY 2316 TCATGACGCTCATCAACACC 2336
DB 2 SerCySerSerSerThrThr 8

RESULT 39
US-10-462-850-2387
/ Sequence 2387, Application US/10462850
/ GENERAL INFORMATION:
/ APPLICANT: Proteom Ltd
/ TITLE OF INVENTION: Complementary peptide ligands from the human genome
/ FILE REFERENCE: Human patent
/ CURRENT APPLICATION NUMBER: US/10/462,850
/ CURRENT FILING DATE: 2003-06-17
/ NUMBER OF SEQ ID NOS: 4203
/ SOFTWARE: ProPatent version 1.0
/ SEQ ID NO 2387
/ LENGTH: 10
/ TYPE: PRF
/ ORGANISM: Homo Sapiens
/ FEATURE:
/ OTHER INFORMATION: sequence located in C8A at 81-90 and may interact with Sequence 2
US-10-462-850-2387

Alignment Scores:
Pred. No.: 117 Length: 10
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-462-850-2387 (1-10)
QY 2316 TCATGACGCTCATCAACACC 2336
DB 3 SerCySerSerSerThrThr 9

RESULT 40
US-10-462-850-2399
/ Sequence 2399, Application US/10462850
/ GENERAL INFORMATION:
/ APPLICANT: Proteom Ltd
/ TITLE OF INVENTION: Complementary peptide ligands from the human genome
/ FILE REFERENCE: Human patent
/ CURRENT APPLICATION NUMBER: US/10/462,850
/ CURRENT FILING DATE: 2003-06-17
/ NUMBER OF SEQ ID NOS: 4203
/ SOFTWARE: ProPatent version 1.0
/ SEQ ID NO 2399
/ LENGTH: 10
```

```

; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 81-90 and may interact with Sequence 2
US-10-462-850-2399
Alignment Scores:
Pred. No.: 117      Length: 10
Score: 7.00      Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.78%      Indels: 0
DB: 6      Gaps: 0
US-09-774-490-1 (1-2709) x US-10-462-850-2399 (1-10)
QY 2316 TCATGACGCTCATCAACCACC 2336
DB 3 SerCysSerSerSerThrThr 9
RESULT 41
US-10-462-850-2401
; Sequence 2401, Application US/10462850
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/10/462,850
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 2401
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 83-92 and may interact with Sequence 2
US-10-462-850-2401
Alignment Scores:
Pred. No.: 117      Length: 10
Score: 7.00      Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.78%      Indels: 0
DB: 6      Gaps: 0
US-09-774-490-1 (1-2709) x US-10-462-850-2401 (1-10)
QY 2316 TCATGACGCTCATCAACCACC 2336
DB 1 SerCysSerSerSerThrThr 7
RESULT 42
US-10-462-850-2485
; Sequence 2485, Application US/10462850
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/10/462,850
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 2485
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 81-90 and may interact with Sequence 2
US-10-462-850-2485

```

```

US-10-462-850-2485
Alignment Scores:
Pred. No.: 117      Length: 10
Score: 7.00      Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.78%      Indels: 0
DB: 6      Gaps: 0
US-09-774-490-1 (1-2709) x US-10-462-850-2485 (1-10)
QY 2316 TCATGACGCTCATCAACCACC 2336
DB 3 SerCysSerSerSerThrThr 9
RESULT 43
US-10-462-850-2527
; Sequence 2527, Application US/10462850
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/10/462,850
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 2527
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 83-92 and may interact with Sequence 2
US-10-462-850-2527
Alignment Scores:
Pred. No.: 117      Length: 10
Score: 7.00      Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.78%      Indels: 0
DB: 6      Gaps: 0
US-09-774-490-1 (1-2709) x US-10-462-850-2527 (1-10)
QY 2316 TCATGACGCTCATCAACCACC 2336
DB 1 SerCysSerSerSerThrThr 7
RESULT 44
US-10-462-850-2533
; Sequence 2533, Application US/10462850
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/10/462,850
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProPatent version 1.0
; SEQ ID NO 2533
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 83-92 and may interact with Sequence 2
US-10-462-850-2533
Alignment Scores:
Pred. No.: 117      Length: 10
Score: 7.00      Matches: 7

```

Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.78%
DB: 6
Gaps: 0

US-09-774-490-1 (1-2709) x US-10-462-850-2533 (1-10)

QY 2316 TCATGACGCTCATCAACACC 2336
|||||
1 SerCysSerSerSerThrThr 7

RESULT 45

US-10-462-850-2543
; Sequence 2543, Application US/10462850
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/10/462,850
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 2543
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 81-90 and may interact with Sequence 2
US-10-462-850-2543

Alignment Scores:

Pred. No.: 117
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.78%
DB: 6
Length: 10
Matches: 7
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-10-462-850-2543 (1-10)

QY 2316 TCATGACGCTCATCAACACC 2336
|||||
3 SerCysSerSerSerThrThr 9

RESULT 46

US-10-462-850-2557
; Sequence 2557, Application US/10462850
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/10/462,850
; CURRENT FILING DATE: 2003-06-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: ProtPatent version 1.0
; SEQ ID NO 2557
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in C8A at 83-92 and may interact with Sequence 2
US-10-462-850-2557

Alignment Scores:

Pred. No.: 117
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.78%
DB: 6
Length: 10
Matches: 7
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-10-462-850-2557 (1-10)

QY 2316 TCATGACGCTCATCAACACC 2336
|||||
1 SerCysSerSerSerThrThr 7

RESULT 47

US-10-601-837-156
; Sequence 156, Application US/10601837
; GENERAL INFORMATION:
; APPLICANT: Holt, Gordon D
; APPLICANT: Kelly, Michael D
; APPLICANT: Kennedy, Sandra J
; APPLICANT: Moyes, Christopher
; TITLE OF INVENTION: Proteins, Genes and Their Use for Diagnosis and Treatment of Kidn
; FILE REFERENCE: 2543-1-030
; CURRENT APPLICATION NUMBER: US/10/601,837
; CURRENT FILING DATE: 2003-06-23
; PRIOR APPLICATION NUMBER: PCT/GB01/05777
; PRIOR FILING DATE: 2001-12-24
; PRIOR APPLICATION NUMBER: US 60/260392
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 272
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 156
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Rattus Norvegicus
US-10-601-837-156

Alignment Scores:

Pred. No.: 115
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.80%
DB: 6
Length: 12
Matches: 7
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-10-601-837-156 (1-12)

QY 1164 GCATTTTGAAGATCTTTAAG 1144
|||||
6 GlyPheLeuGlySerLeuLys 12

RESULT 48

PCT-US02-39397-69
; Sequence 69, Application PC/TUS0239397
; GENERAL INFORMATION:
; APPLICANT: Iowa State University Research Foundation, Inc.
; TITLE OF INVENTION: NUCLEIC ACIDS RELATED TO PLANT
; TITLE OF INVENTION: RETROELEMENTS
; FILE REFERENCE: 08411-031M01
; CURRENT APPLICATION NUMBER: PCT/US02/39397
; CURRENT FILING DATE: 2003-06-10
; PRIOR APPLICATION NUMBER: US 60/339,060
; PRIOR FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 168
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 69
; LENGTH: 22
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
PCT-US02-39397-69

Alignment Scores:

Pred. No.: 110
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.80%
DB: 1
Length: 22
Matches: 7
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US02-39397-69 (1-22)

Qy 2475 CTGTTCTACCTTCTTATTT 2455
Db 16 Leupheleupheleuphe 22

RESULT 49

US-10-369-072-90

/ Sequence 90, Application US/10369072

/ GENERAL INFORMATION:

/ APPLICANT: Alsobrook II, John P

/ APPLICANT: Spaderma, Stephen K

/ APPLICANT: Tchernev, Velizar

/ APPLICANT: Liu, Xiaohong

/ APPLICANT: Shenoy, Suresh

/ APPLICANT: Spytek, Kimberly

/ APPLICANT: Zernusen, Bryan

/ APPLICANT: Paturajan, Weera

/ APPLICANT: Taupier, Raymond T

/ APPLICANT: Rastelli, Luca

/ APPLICANT: Grosse, William M

/ APPLICANT: Szerkeres, Edward S

/ APPLICANT: Lepley, Denise M

/ APPLICANT: Shen, Lei

/ APPLICANT: Burgess, Catherine E

/ APPLICANT: Shinkets, Richard

/ APPLICANT: Padigaru, Muralidhara

/ TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same

/ FILE REFERENCE: 21402-050 CON2

/ CURRENT APPLICATION NUMBER: US/10/369,072

/ PRIOR FILING DATE: 2003-02-18

/ PRIOR APPLICATION NUMBER: 10/174,372

/ PRIOR FILING DATE: 2002-06-17

/ PRIOR APPLICATION NUMBER: 09/898,994

/ PRIOR FILING DATE: 2001-07-03

/ PRIOR APPLICATION NUMBER: 60/215,854

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR APPLICATION NUMBER: 60/215,856

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR APPLICATION NUMBER: 60/215,902

/ PRIOR FILING DATE: 2000-07-03

/ PRIOR APPLICATION NUMBER: 60/216,585

/ PRIOR FILING DATE: 2000-07-07

/ PRIOR APPLICATION NUMBER: 60/216,586

/ PRIOR FILING DATE: 2000-07-07

/ PRIOR APPLICATION NUMBER: 60/216,722

/ PRIOR FILING DATE: 2000-07-07

/ PRIOR APPLICATION NUMBER: 60/218,622

/ PRIOR FILING DATE: 2000-07-17

/ PRIOR APPLICATION NUMBER: 60/218,992

/ PRIOR FILING DATE: 2000-07-17

/ Remaining Prior Application data removed - See File Wrapper or PALM.

/ NUMBER OF SEQ ID NOS: 100

/ SOFTWARE: PatentIn Ver. 2.1

/ SEQ ID NO 90

/ LENGTH: 47

/ TYPE: PRT

/ ORGANISM: Artificial Sequence

/ FEATURE:

/ OTHER INFORMATION: Description of Artificial Sequence: Domain search

/ US-10-369-072-90

Alignment Scores:

Pred. No.: 104

Score: 7.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 0.78%

DB: 6

US-09-774-490-1 (1-2709) x US-10-369-072-90 (1-47)

Qy 1784 CTCGCCGAGACCTTACTGT 1804
Db 13 LeuAlaArgAspProTyrCys 19

RESULT 50

US-10-603-114-7994

/ Sequence 7994, Application US/10603114

/ GENERAL INFORMATION:

/ APPLICANT: GARY BRETON

/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

/ FILE REFERENCE: 2709.1002-001

/ CURRENT APPLICATION NUMBER: US/10/603,114

/ PRIOR FILING DATE: 2003-06-24

/ PRIOR APPLICATION NUMBER: US/09/543,681

/ PRIOR FILING DATE: 2000-04-05

/ PRIOR APPLICATION NUMBER: US 60/128,706

/ PRIOR FILING DATE: 1999-04-09

/ NUMBER OF SEQ ID NOS: 8344

/ SEQ ID NO 7994

/ LENGTH: 72

/ TYPE: PRT

/ ORGANISM: Proteus mirabilis

/ US-10-603-114-7994

Alignment Scores:

Pred. No.: 100

Score: 7.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 0.80%

DB: 6

US-09-774-490-1 (1-2709) x US-10-603-114-7994 (1-72)

Qy 2308 GTACCGACCTTGTGCTAGG 2288

Db 42 ValProAspLeuAlaIleArg 48

RESULT 51

US-10-374-780A-2152

/ Sequence 2152, Application US/10374780A

/ GENERAL INFORMATION:

/ APPLICANT: Sherman, Bradley K

/ APPLICANT: Riechmann, Jose Luis

/ APPLICANT: Jiang, Cai-Zhong

/ APPLICANT: Heard, Jacqueline E

/ APPLICANT: Haake, Volker

/ APPLICANT: Creelman, Robert A

/ APPLICANT: Ratcliffe, Oliver

/ APPLICANT: Adam, Luc J

/ APPLICANT: Reuber, T. Lynne

/ APPLICANT: Keddie, James

/ APPLICANT: Brown, Pierre E

/ APPLICANT: Pilgrim, Marsha L

/ APPLICANT: Dubell III, Arnold T

/ APPLICANT: Pineda, Omeira

/ APPLICANT: Yu, Guo-Liang

/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS

/ FILE REFERENCE: MBI-0047 CIP

/ CURRENT APPLICATION NUMBER: US/10/374,780A

/ PRIOR FILING DATE: 2003-02-25

/ PRIOR APPLICATION NUMBER: 09/837,944

/ PRIOR FILING DATE: 2001-04-18

/ PRIOR APPLICATION NUMBER: 60/310,847

/ PRIOR FILING DATE: 2001-08-09

/ PRIOR APPLICATION NUMBER: 09/934,455

/ PRIOR FILING DATE: 2001-08-22

/ PRIOR APPLICATION NUMBER: 60/336,049

/ PRIOR FILING DATE: 2001-11-19

/ PRIOR APPLICATION NUMBER: 60/338,692

/ PRIOR FILING DATE: 2001-12-11

/ PRIOR APPLICATION NUMBER: 10/171,468

```
/ PRIOR FILING DATE: 2002-06-14
/ PRIOR APPLICATION NUMBER: 10/225,066
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: 10/225,067
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: 10/225,068
/ PRIOR FILING DATE: 2002-08-09
/ NUMBER OF SEQ ID NOS: 2906
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 2152
/ LENGTH: 121
/ TYPE: PRT
/ ORGANISM: Arabidopsis thaliana
/ FEATURE:
/ OTHER INFORMATION: G1929 Paralogous to G1478
US-10-374-780A-2152

Alignment Scores:
Pred. No.: 96.6 Length: 121
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-374-780A-2152 (1-121)

QY 2315 AGCTCTGTACCGAGCTTCT 2295
DB 20 SerLeuCystrHrargProser 26

RESULT 52
US-10-273-573-7304
/ Sequence 7304, Application US/10273573
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc
/ TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
/ FILE REFERENCE: 21272-066
/ CURRENT APPLICATION NUMBER: US/10/273,573
/ PRIOR FILING DATE: 2002-10-18
/ PRIOR APPLICATION NUMBER: 09/522,929
/ PRIOR FILING DATE: 2000-04-18
/ PRIOR APPLICATION NUMBER: 09/770,160
/ PRIOR FILING DATE: 2001-01-26
/ NUMBER OF SEQ ID NOS: 10994
/ SOFTWARE: Custom
/ SEQ ID NO 7304
/ LENGTH: 126
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: DOMAIN
/ LOCATION: (109)..(125)
/ OTHER INFORMATION: C4-TYPE STEROID RECEPTOR ZINC FINGER SIGNATURE domain
/ OTHER INFORMATION: identified by eMBL, accession number PR00047B, p-value=1.000e-
/ OTHER INFORMATION: 20, raw score of 7.63
/ FEATURE:
/ NAME/KEY: DOMAIN
/ LOCATION: (110)..(126)
/ OTHER INFORMATION: Zinc finger, C4 type (two domains) domain identified by Pfam,
/ OTHER INFORMATION: accession name zf-C4, E-value=5.1e-09, Pfam score of 34.4
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(126)
/ OTHER INFORMATION: Xaa = X or * as defined in Table 2
US-10-273-573-7304

Alignment Scores:
Pred. No.: 96.3 Length: 126
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
```

```
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-273-573-7304 (1-126)

QY 1024 CCTCCAAAGTCATTCTTGA 1004
DB 62 ProserLySvalIleuVala 68

RESULT 53
US-10-374-780A-1248
/ Sequence 1248, Application US/10374780A
/ GENERAL INFORMATION:
/ APPLICANT: Sherman, Bradley K
/ APPLICANT: Riechmann, Jose Luis
/ APPLICANT: Jiang, Cai-Zhong
/ APPLICANT: Heard, Jacqueline E
/ APPLICANT: Haake, Volker
/ APPLICANT: Creelman, Robert A
/ APPLICANT: Ratcliffe, Oliver
/ APPLICANT: Adam, Luc J
/ APPLICANT: Reuber, T. Lynne
/ APPLICANT: Keddie, James
/ APPLICANT: Brown, Pierre E
/ APPLICANT: Pilgrim, Marsha L
/ APPLICANT: Dubeil III, Arnold T
/ APPLICANT: Pineda, Omaira
/ APPLICANT: Yu, Guo-Liang
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
/ FILE REFERENCE: MBI-0047 CIP
/ CURRENT APPLICATION NUMBER: US/10/374,780A
/ PRIOR FILING DATE: 2003-02-25
/ PRIOR APPLICATION NUMBER: 09/837,944
/ PRIOR FILING DATE: 2001-04-18
/ PRIOR APPLICATION NUMBER: 60/310,847
/ PRIOR FILING DATE: 2001-08-09
/ PRIOR APPLICATION NUMBER: 09/934,455
/ PRIOR FILING DATE: 2001-08-22
/ PRIOR APPLICATION NUMBER: 60/336,049
/ PRIOR FILING DATE: 2001-11-19
/ PRIOR APPLICATION NUMBER: 60/338,692
/ PRIOR FILING DATE: 2001-12-11
/ PRIOR APPLICATION NUMBER: 10/171,468
/ PRIOR FILING DATE: 2002-06-14
/ PRIOR APPLICATION NUMBER: 10/225,066
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: 10/225,067
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: 10/225,068
/ PRIOR FILING DATE: 2002-08-09
/ NUMBER OF SEQ ID NOS: 2906
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 1248
/ LENGTH: 145
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Orthologous to G912
US-10-374-780A-1248

Alignment Scores:
Pred. No.: 95.3 Length: 145
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-374-780A-1248 (1-145)

QY 1615 AGAAGAGTTCTGCTGGAAGA 1635
DB 60 ArgArgGlySerAlaGlyArg 66
```

RESULT 54
US-10-374-780A-1342
; Sequence 1342, Application US/10374780A
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keadie, James
; APPLICANT: Brown, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MBI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1342
; LENGTH: 151
; TYPE: PRT
; ORGANISM: *Oryza sativa*
; FEATURE:
; OTHER INFORMATION: Orthologous to G1047
US-10-374-780A-1342

Alignment Scores:
Pred. No.: 95 Length: 151
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-374-780A-1342 (1-151)
QY 2253 |||||
DB 28 ProserProserSerieu 34

RESULT 55
US-10-603-114-7828
; Sequence 7828, Application US/10603114
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001

; CURRENT APPLICATION NUMBER: US/10/603,114
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/543,681
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 7828
; LENGTH: 192
; TYPE: PRT
; ORGANISM: *Proteus mirabilis*
US-10-603-114-7828

Alignment Scores:
Pred. No.: 93.3 Length: 192
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-603-114-7828 (1-192)
QY 1403 |||||
DB 103 PhenlaArgSerHisProAla 109

RESULT 56
US-10-603-114-6794
; Sequence 6794, Application US/10603114
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/10/603,114
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/543,681
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6794
; LENGTH: 193
; TYPE: PRT
; ORGANISM: *Proteus mirabilis*
US-10-603-114-6794

Alignment Scores:
Pred. No.: 93.2 Length: 193
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-603-114-6794 (1-193)
QY 27 AGCTTAGTATGATGCCACG 47
DB 131 SerleuValIleleapAlaThr 137

RESULT 57
US-10-293-244-3807
; Sequence 3807, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hyseng, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned

```
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 09/728,422
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: 09/693,325
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 09/663,561
/ PRIOR FILING DATE: 2000-09-15
/ PRIOR APPLICATION NUMBER: 09/654,936
/ PRIOR FILING DATE: 2000-09-01
/ PRIOR APPLICATION NUMBER: 09/620,325
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/598,075
/ PRIOR FILING DATE: 2000-06-20
/ PRIOR APPLICATION NUMBER: 09/560,875
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: 09/496,914
/ PRIOR FILING DATE: 2000-02-03
/ NUMBER OF SEQ ID NOS: 3960
/ SOFTWARE: Custom
/ SEQ ID NO: 3807
/ LENGTH: 194
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(194)
/ OTHER INFORMATION: Xaa = any amino acid or nothing
US-10-293-244-3807
```

```
Alignment Scores:
Pred. No.: 93.2 Length: 194
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 6 Gaps: 0
```

US-09-774-490-1 (1-2709) x US-10-293-244-3807 (1-194)

QY 2413 TCCTGCGCTTGGCGGCTTG 2393

Db 9 Serritroleuphrothleu 15

```
RESULT 58
US-10-374-780A-414
/ Sequence 414, Application US/10374780A
/ GENERAL INFORMATION:
/ APPLICANT: Sherman, Bradley K
/ APPLICANT: Riechmann, Jose Luis
/ APPLICANT: Jiang, Cai-Zhong
/ APPLICANT: Heard, Jacqueline E
/ APPLICANT: Haake, Volker
/ APPLICANT: Creelman, Robert A
/ APPLICANT: Ratcliffe, Oliver
/ APPLICANT: Adam, Luc J
/ APPLICANT: Reuber, T. Lynne
/ APPLICANT: Keddie, James
/ APPLICANT: Brown, Pierre E
/ APPLICANT: Pilgrim, Marsha L
/ APPLICANT: Dubell III, Arnold T
/ APPLICANT: Pineda, Omaliza
/ APPLICANT: Yu, Guo-Liang
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
/ FILE REFERENCE: MBI-0047 CIP
/ CURRENT APPLICATION NUMBER: US/10/374,780A
/ PRIOR FILING DATE: 2003-02-25
/ PRIOR APPLICATION NUMBER: 09/837,944
/ PRIOR FILING DATE: 2001-04-18
/ PRIOR APPLICATION NUMBER: 60/310,847
/ PRIOR FILING DATE: 2001-08-09
/ PRIOR APPLICATION NUMBER: 09/934,455
/ PRIOR FILING DATE: 2001-08-22
/ PRIOR APPLICATION NUMBER: 60/336,049
```

```
/ PRIOR FILING DATE: 2001-11-19
/ PRIOR APPLICATION NUMBER: 60/338,692
/ PRIOR FILING DATE: 2001-12-11
/ PRIOR APPLICATION NUMBER: 10/171,468
/ PRIOR FILING DATE: 2002-06-14
/ PRIOR APPLICATION NUMBER: 10/225,066
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: 10/225,067
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: 10/225,068
/ PRIOR FILING DATE: 2002-08-09
/ NUMBER OF SEQ ID NOS: 2906
/ SOFTWARE: Patent in version 3.2
/ SEQ ID NO: 414
/ LENGTH: 231
/ TYPE: PRT
/ ORGANISM: Arabidopsis thaliana
/ FEATURE:
/ OTHER INFORMATION: G2143 (conserved domain in aa coordinates: 128-179)
US-10-374-780A-414
```

```
Alignment Scores:
Pred. No.: 92 Length: 231
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 6 Gaps: 0
```

US-09-774-490-1 (1-2709) x US-10-374-780A-414 (1-231)

QY 2253 CCATCTCATCATCATCTTTA 2233

Db 73 ProserProserSerleu 79

```
RESULT 59
US-10-374-780A-1246
/ Sequence 1246, Application US/10374780A
/ GENERAL INFORMATION:
/ APPLICANT: Sherman, Bradley K
/ APPLICANT: Riechmann, Jose Luis
/ APPLICANT: Jiang, Cai-Zhong
/ APPLICANT: Heard, Jacqueline E
/ APPLICANT: Haake, Volker
/ APPLICANT: Creelman, Robert A
/ APPLICANT: Ratcliffe, Oliver
/ APPLICANT: Adam, Luc J
/ APPLICANT: Reuber, T. Lynne
/ APPLICANT: Keddie, James
/ APPLICANT: Brown, Pierre E
/ APPLICANT: Pilgrim, Marsha L
/ APPLICANT: Dubell III, Arnold T
/ APPLICANT: Pineda, Omaliza
/ APPLICANT: Yu, Guo-Liang
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
/ FILE REFERENCE: MBI-0047 CIP
/ CURRENT APPLICATION NUMBER: US/10/374,780A
/ PRIOR FILING DATE: 2003-02-25
/ PRIOR APPLICATION NUMBER: 09/837,944
/ PRIOR FILING DATE: 2001-04-18
/ PRIOR APPLICATION NUMBER: 60/310,847
/ PRIOR FILING DATE: 2001-08-09
/ PRIOR APPLICATION NUMBER: 09/934,455
/ PRIOR FILING DATE: 2001-08-22
/ PRIOR APPLICATION NUMBER: 60/336,049
/ PRIOR FILING DATE: 2001-11-19
/ PRIOR APPLICATION NUMBER: 60/338,692
/ PRIOR FILING DATE: 2001-12-11
/ PRIOR APPLICATION NUMBER: 10/171,468
/ PRIOR FILING DATE: 2002-06-14
/ PRIOR APPLICATION NUMBER: 10/225,066
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: 10/225,067
```



```
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: 10/225,068
/ PRIOR FILING DATE: 2002-08-09
/ NUMBER OF SEQ ID NOS: 2906
/ SOFTWARE: PatentIn version 3.2
/ SEQ ID NO 1246
/ LENGTH: 231
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Orthologous to G912
US-10-374-780A-1246

Alignment Scores:
Pred. No.: 92          Length: 231
Score: 7.00           Matches: 7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 0.78%          Indels: 0
DB: 6                  Gaps: 0

US-09-774-490-1 (1-2709) x US-10-374-780A-1246 (1-231)
QY      1615 AGAAGAGTCTCTGGAAGA 1635
Db      58 ArgArgGlySerAlaGlyArg 64

RESULT 60
US-10-603-114-4250
/ Sequence 4250, Application US/10603114
/ GENERAL INFORMATION:
/ APPLICANT: GARY BRETON
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
/ FILE REFERENCE: 2709.1002-001
/ CURRENT APPLICATION NUMBER: US/10/603,114
/ PRIOR FILING DATE: 2003-06-24
/ PRIOR APPLICATION NUMBER: US/09/543,681
/ PRIOR FILING DATE: 2000-04-05
/ PRIOR APPLICATION NUMBER: US 60/128,706
/ PRIOR FILING DATE: 1999-04-09
/ NUMBER OF SEQ ID NOS: 8344
/ SEQ ID NO 4250
/ LENGTH: 251
/ TYPE: PRT
/ ORGANISM: Proteus mirabilis
US-10-603-114-4250

Alignment Scores:
Pred. No.: 91.4        Length: 251
Score: 7.00           Matches: 7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 0.80%          Indels: 0
DB: 6                  Gaps: 0

US-09-774-490-1 (1-2709) x US-10-603-114-4250 (1-251)
QY      1699 TTGTTGCTGCTTAGTGAAG 1679
Db      177 LeuLeuLeuSerGlyLys 183

RESULT 61
US-10-293-244-3649
/ Sequence 3649, Application US/10293244
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc.
/ APPLICANT: Tang, Y. Tom et al
/ TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
/ FILE REFERENCE: 21272-029
/ CURRENT APPLICATION NUMBER: US/10/293,244
/ CURRENT FILING DATE: 2002-11-12
/ PRIOR APPLICATION NUMBER: Not yet Assigned
```

```
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 09/728,422
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: 09/693,325
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 09/663,561
/ PRIOR FILING DATE: 2000-09-15
/ PRIOR APPLICATION NUMBER: 09/654,936
/ PRIOR FILING DATE: 2000-09-01
/ PRIOR APPLICATION NUMBER: 09/620,325
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/598,075
/ PRIOR FILING DATE: 2000-06-20
/ PRIOR APPLICATION NUMBER: 09/560,875
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: 09/496,914
/ PRIOR FILING DATE: 2000-02-03
/ NUMBER OF SEQ ID NOS: 3960
/ SOFTWARE: Custom
/ SEQ ID NO 3649
/ LENGTH: 258
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-293-244-3649

Alignment Scores:
Pred. No.: 91.2        Length: 258
Score: 7.00           Matches: 7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 0.78%          Indels: 0
DB: 6                  Gaps: 0

US-09-774-490-1 (1-2709) x US-10-293-244-3649 (1-258)
QY      728 ACAGCATCCCTTTATAGAT 748
Db      142 ThrAlaSerLeuLeuIleAsp 148

RESULT 62
US-10-603-114-5243
/ Sequence 5243, Application US/10603114
/ GENERAL INFORMATION:
/ APPLICANT: GARY BRETON
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
/ FILE REFERENCE: 2709.1002-001
/ CURRENT APPLICATION NUMBER: US/10/603,114
/ PRIOR FILING DATE: 2003-06-24
/ PRIOR APPLICATION NUMBER: US/09/543,681
/ PRIOR FILING DATE: 2000-04-05
/ PRIOR APPLICATION NUMBER: US 60/128,706
/ PRIOR FILING DATE: 1999-04-09
/ NUMBER OF SEQ ID NOS: 8344
/ SEQ ID NO 5243
/ LENGTH: 269
/ TYPE: PRT
/ ORGANISM: Proteus mirabilis
US-10-603-114-5243

Alignment Scores:
Pred. No.: 90.9        Length: 269
Score: 7.00           Matches: 7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 0.78%          Indels: 0
DB: 6                  Gaps: 0

US-09-774-490-1 (1-2709) x US-10-603-114-5243 (1-269)
QY      1593 CTAAGAGACTTGGTATGATT 1613
Db      188 LeuArgGlyLeuGlyMetIle 194
```

```
RESULT 63
US-10-292-798-1056
; Sequence 1056, Application US/10292798
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 08435/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1056
; LENGTH: 271
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-292-798-1056

Alignment Scores:
Pred. No.: 90.9      Length: 271
Score: 7.00         Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.78%              Indels: 0
DB: 6                  Gaps: 0

US-09-774-490-1 (1-2709) x US-10-292-798-1056 (1-271)

QY 1160 AATCCAGTTGTATATGAGTGG 1180
DB 240 AsnProValValtyrGlyVal 246

RESULT 64
US-10-603-114-5282
; Sequence 5282, Application US/10603114
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/10/603,114
; CURRENT FILING DATE: 2003-06-24
; PRIOR APPLICATION NUMBER: US/09/543,681
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 5282
; LENGTH: 287
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-10-603-114-5282

Alignment Scores:
Pred. No.: 90.5      Length: 287
Score: 7.00         Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.78%              Indels: 0
DB: 6                  Gaps: 0

US-09-774-490-1 (1-2709) x US-10-603-114-5282 (1-287)

QY 1644 TTTTGGGAGACCGACTGCTA 1664
DB 275 PhepeglyAsnArgLeuLeu 281
```

```
RESULT 65
US-09-665-308E-2
; Sequence 2, Application US/09665308E
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Klein, Theodore M.
; APPLICANT: Odell, Joan T.
; APPLICANT: Orozco, Emil M. Jr.
; TITLE OF INVENTION: PLANT CELL CYCLIN GENES
; FILE REFERENCE: B81149 US NA
; CURRENT APPLICATION NUMBER: US/09/665,308E
; CURRENT FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/078,735
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: PCT/US99/06047
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 32
; SOFTWARE: MICROSOFT OFFICE 97
; SEQ ID NO 2
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Zea mays
US-09-665-308E-2

Alignment Scores:
Pred. No.: 90.3      Length: 295
Score: 7.00         Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.78%              Indels: 0
DB: 5                  Gaps: 0

US-09-774-490-1 (1-2709) x US-09-665-308E-2 (1-295)

QY 54 GGTTGACCCACGCGTCCGG 74
DB 4 GlyValAspProArgValArg 10

RESULT 66
US-10-374-780A-1658
; Sequence 1658, Application US/10374780A
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jiang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddle, James
; APPLICANT: Brown, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MB1-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; CURRENT FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
```

;; PRIOR FILING DATE: 2002-06-14
;; PRIOR APPLICATION NUMBER: 10/225,066
;; PRIOR FILING DATE: 2002-08-09
;; PRIOR APPLICATION NUMBER: 10/225,067
;; PRIOR FILING DATE: 2002-08-09
;; PRIOR APPLICATION NUMBER: 10/225,068
;; PRIOR FILING DATE: 2002-08-09
;; NUMBER OF SEQ ID NOS: 2906
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO: 1658
;; LENGTH: 303
;; TYPE: PRT
;; ORGANISM: Oryza sativa
;; FEATURE:
;; OTHER INFORMATION: Orthologous to G1499
US-10-374-780A-1658

Alignment Scores:
Pred. No.: 90.1 Length: 303
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-374-780A-1658 (1-303)

QY 2253 CCATCGATCATCATCTTA 2233

DB 52 ProSerProSerSerSer 58

RESULT 67

US-10-610-917-5
;; Sequence 5, Application US/10610917
;; GENERAL INFORMATION:
;; APPLICANT: Olsen et al.
;; TITLE OF INVENTION: Human Freezied-Like Protein
;; FILE REFERENCE: PF389C1
;; CURRENT APPLICATION NUMBER: US/10/610,917
;; CURRENT FILING DATE: 2003-07-02
;; PRIOR APPLICATION NUMBER: US/09/610,650B
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: 09/132,315
;; PRIOR FILING DATE: 1998-08-11
;; PRIOR APPLICATION NUMBER: 60/055,715
;; PRIOR FILING DATE: 1997-08-12
;; PRIOR APPLICATION NUMBER: 60/081,438
;; PRIOR FILING DATE: 1998-04-10
;; NUMBER OF SEQ ID NOS: 27
;; SOFTWARE: PatentIn version 3.1
;; SEQ ID NO: 5
;; LENGTH: 325
;; TYPE: PRT
;; ORGANISM: Bos taurus
US-10-610-917-5

Alignment Scores:
Pred. No.: 89.6 Length: 325
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-610-917-5 (1-325)

QY 386 GATGAGAAAGGATAGGCTG 406

DB 261 AepgluGluaArgSerArgLeu 267

RESULT 68

US-10-292-798-798
;; Sequence 798, Application US/10292798

;; GENERAL INFORMATION:
;; APPLICANT: SUMA, MAKIRO
;; APPLICANT: ASAI, KIYOSHI
;; APPLICANT: AKIYAMA, YUTAKA
;; APPLICANT: ABEHATANI, HIROYUKI
;; TITLE OF INVENTION: GANGLIOSINE TIPOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
;; FILE REFERENCE: 084335/166
;; CURRENT APPLICATION NUMBER: US/10/292,798
;; CURRENT FILING DATE: 2002-11-13
;; PRIOR APPLICATION NUMBER: 10/017,161
;; PRIOR FILING DATE: 2001-12-18
;; PRIOR APPLICATION NUMBER: JP 2001-246789
;; PRIOR FILING DATE: 2001-06-18
;; NUMBER OF SEQ ID NOS: 2070
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO: 798
;; LENGTH: 329
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-292-798-798

Alignment Scores:
Pred. No.: 89.5 Length: 329
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-292-798-798 (1-329)

QY 1160 AATCAGTTGATATGAGTG 1180

DB 298 AsnProValValIYgIYVal 304

RESULT 69

US-10-293-244-1681
;; Sequence 1681, Application US/10293244
;; GENERAL INFORMATION:
;; APPLICANT: Hyseq, Inc.
;; APPLICANT: Tang, Y. Tom et al
;; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
;; FILE REFERENCE: 21272-029
;; CURRENT APPLICATION NUMBER: US/10/293,244
;; CURRENT FILING DATE: 2002-11-12
;; PRIOR APPLICATION NUMBER: Not Yet Assigned
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: 09/728,422
;; PRIOR FILING DATE: 2000-11-30
;; PRIOR APPLICATION NUMBER: 09/693,325
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 09/663,561
;; PRIOR FILING DATE: 2000-09-15
;; PRIOR APPLICATION NUMBER: 09/654,936
;; PRIOR FILING DATE: 2000-09-01
;; PRIOR APPLICATION NUMBER: 09/620,325
;; PRIOR FILING DATE: 2000-07-19
;; PRIOR APPLICATION NUMBER: 09/598,075
;; PRIOR FILING DATE: 2000-06-20
;; PRIOR APPLICATION NUMBER: 09/560,875
;; PRIOR FILING DATE: 2000-04-27
;; PRIOR APPLICATION NUMBER: 09/496,914
;; PRIOR FILING DATE: 2000-02-03
;; NUMBER OF SEQ ID NOS: 3960
;; SOFTWARE: Custom
;; SEQ ID NO: 1681
;; LENGTH: 332
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-293-244-1681

Alignment Scores:
Pred. No.: 89.5 Length: 332

Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-293-244-1681 (1-332)

QY 728 ACAGCATCCCTTTATAGAT 748

DB 145 ThrAlaSerLeuLeuLeuLeu 151

RESULT 70

PCT-US03-21576-38
Sequence 38, Application PC/TUS0321576

GENERAL INFORMATION:

APPLICANT: INCYTE CORPORATION

APPLICANT: KABLE, Amy E.

APPLICANT: SWARNARKAR, Anita

APPLICANT: LEE, Soo Yeun

APPLICANT: RICHARDSON, Thomas W.

APPLICANT: ISON, Craig H.

APPLICANT: KHARE, Reena

APPLICANT: JIN, Pei

APPLICANT: FU, Glenn K.

APPLICANT: BECHA, Shanya D.

APPLICANT: WANG, Jonathan T.

APPLICANT: TANG, Y. Tom

APPLICANT: ELLIOTT, Vicki S.

APPLICANT: WARREN, Bridget A.

APPLICANT: JIANG, Xin

APPLICANT: JACKSON, Alan A.

APPLICANT: BAUGHN, Mariah R.

APPLICANT: DUGAN, Brendan M.

APPLICANT: HAPALIA, April J.A.

APPLICANT: RAMKUMAR, Jayalaxmi

APPLICANT: TRAN, Uyen K.

APPLICANT: LINDQUIST, Erika A.

APPLICANT: CHAWLA, Nandinder K.

APPLICANT: MARQUIS, Joseph P.

TITLE OF INVENTION: SECRETED PROTEINS

FILE REFERENCE: PF-1476 PCT

CURRENT APPLICATION NUMBER: PCT/US03/21576

PRIOR FILING DATE: 2003-07-10

PRIOR APPLICATION NUMBER: US 60/395,722

PRIOR FILING DATE: 2002-07-11

PRIOR APPLICATION NUMBER: US 60/399,003

PRIOR FILING DATE: 2002-07-26

PRIOR APPLICATION NUMBER: US 60/400,909

PRIOR FILING DATE: 2002-08-01

PRIOR APPLICATION NUMBER: US 60/407,552

PRIOR FILING DATE: 2002-08-30

PRIOR APPLICATION NUMBER: US 60/407,202

PRIOR FILING DATE: 2002-08-30

NUMBER OF SEQ ID NOS: 108

SOFTWARE: PERL Program

SEO ID NO 38

LENGTH: 333

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No: 7513709CD1

PCT-US03-21576-38

Alignment Scores:

Pred. No.: 89.5

Score: 7.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 0.78%

US-09-774-490-1 (1-2709) x PCT-US03-21576-38 (1-333)

QY 722 CTGCTGACGATCCCTTTA 742

DB 19 LeuLeuThrAlaSerLeuLeu 25

RESULT 71

PCT-US03-21576-39

Sequence 39, Application PC/TUS0321576

GENERAL INFORMATION:

APPLICANT: INCYTE CORPORATION

APPLICANT: KABLE, Amy E.

APPLICANT: SWARNARKAR, Anita

APPLICANT: LEE, Soo Yeun

APPLICANT: RICHARDSON, Thomas W.

APPLICANT: ISON, Craig H.

APPLICANT: KHARE, Reena

APPLICANT: JIN, Pei

APPLICANT: FU, Glenn K.

APPLICANT: BECHA, Shanya D.

APPLICANT: WANG, Jonathan T.

APPLICANT: TANG, Y. Tom

APPLICANT: ELLIOTT, Vicki S.

APPLICANT: WARREN, Bridget A.

APPLICANT: JIANG, Xin

APPLICANT: JACKSON, Alan A.

APPLICANT: BAUGHN, Mariah R.

APPLICANT: DUGAN, Brendan M.

APPLICANT: HAPALIA, April J.A.

APPLICANT: RAMKUMAR, Jayalaxmi

APPLICANT: TRAN, Uyen K.

APPLICANT: LINDQUIST, Erika A.

APPLICANT: CHAWLA, Nandinder K.

APPLICANT: MARQUIS, Joseph P.

TITLE OF INVENTION: SECRETED PROTEINS

FILE REFERENCE: PF-1476 PCT

CURRENT APPLICATION NUMBER: PCT/US03/21576

PRIOR FILING DATE: 2003-07-10

PRIOR APPLICATION NUMBER: US 60/395,722

PRIOR FILING DATE: 2002-07-11

PRIOR APPLICATION NUMBER: US 60/399,003

PRIOR FILING DATE: 2002-07-26

PRIOR APPLICATION NUMBER: US 60/400,909

PRIOR FILING DATE: 2002-08-01

PRIOR APPLICATION NUMBER: US 60/407,552

PRIOR FILING DATE: 2002-08-30

PRIOR APPLICATION NUMBER: US 60/407,202

PRIOR FILING DATE: 2002-08-30

NUMBER OF SEQ ID NOS: 108

SOFTWARE: PERL Program

SEO ID NO 39

LENGTH: 333

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No: 7513712CD1

PCT-US03-21576-39

Alignment Scores:

Pred. No.: 89.5

Score: 7.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 0.78%

US-09-774-490-1 (1-2709) x PCT-US03-21576-39 (1-333)

QY 722 CTGCTGACGATCCCTTTA 742

DB 19 LeuLeuThrAlaSerLeuLeu 25

RESULT 71

PCT-US03-21576-39

Sequence 39, Application PC/TUS0321576

GENERAL INFORMATION:

APPLICANT: INCYTE CORPORATION

APPLICANT: KABLE, Amy E.

APPLICANT: SWARNARKAR, Anita

APPLICANT: LEE, Soo Yeun

APPLICANT: RICHARDSON, Thomas W.

APPLICANT: ISON, Craig H.

APPLICANT: KHARE, Reena

APPLICANT: JIN, Pei

APPLICANT: FU, Glenn K.

APPLICANT: BECHA, Shanya D.

APPLICANT: WANG, Jonathan T.

APPLICANT: TANG, Y. Tom

APPLICANT: ELLIOTT, Vicki S.

APPLICANT: WARREN, Bridget A.

APPLICANT: JIANG, Xin

APPLICANT: JACKSON, Alan A.

APPLICANT: BAUGHN, Mariah R.

APPLICANT: DUGAN, Brendan M.

APPLICANT: HAPALIA, April J.A.

APPLICANT: RAMKUMAR, Jayalaxmi

APPLICANT: TRAN, Uyen K.

APPLICANT: LINDQUIST, Erika A.

APPLICANT: CHAWLA, Nandinder K.

APPLICANT: MARQUIS, Joseph P.

TITLE OF INVENTION: SECRETED PROTEINS

FILE REFERENCE: PF-1476 PCT

CURRENT APPLICATION NUMBER: PCT/US03/21576

PRIOR FILING DATE: 2003-07-10

PRIOR APPLICATION NUMBER: US 60/395,722

PRIOR FILING DATE: 2002-07-11

PRIOR APPLICATION NUMBER: US 60/399,003

PRIOR FILING DATE: 2002-07-26

PRIOR APPLICATION NUMBER: US 60/400,909

PRIOR FILING DATE: 2002-08-01

PRIOR APPLICATION NUMBER: US 60/407,552

PRIOR FILING DATE: 2002-08-30

PRIOR APPLICATION NUMBER: US 60/407,202

PRIOR FILING DATE: 2002-08-30

NUMBER OF SEQ ID NOS: 108

SOFTWARE: PERL Program

SEO ID NO 39

LENGTH: 333

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: misc feature

OTHER INFORMATION: Incyte ID No: 7513712CD1

PCT-US03-21576-39

Alignment Scores:

Pred. No.: 89.5

Score: 7.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 0.78%

```
RESULT 72
PCT-US03-21703-13
; Sequence 13, Application PC/TUS0321703
; GENERAL INFORMATION:
; APPLICANT: Rupp, Fabio
; APPLICANT: Wang, Jiantui
; APPLICANT: Zhou, Ping
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Zhiwei
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Methods and Materials Relating to Novel Polypeptides and Polynuc
; FILE REFERENCE: HYS-B2 PCT
; CURRENT APPLICATION NUMBER: PCT/US03/21703
; PRIOR FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: PCT/US02/29636
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/395,402
; PRIOR FILING DATE: 2002-07-12
; PRIOR APPLICATION NUMBER: PCT/US02/22858
; PRIOR FILING DATE: 2002-07-19
; PRIOR APPLICATION NUMBER: 10/112,944
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: 60/306,971
; PRIOR FILING DATE: 2001-07-21
; PRIOR APPLICATION NUMBER: 10/296,115
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: PCT/US00/35017
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 60/418,132
; PRIOR FILING DATE: 2002-10-11
; PRIOR APPLICATION NUMBER: PCT/US02/39555
; PRIOR FILING DATE: 2002-12-10
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 13
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Mus musculus
PCT-US03-21703-13

Alignment Scores:
Pred. No.: 89.3      Length: 341
Score: 7.00         Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78%   Indels: 0
DB: 1               Gaps: 0

US-09-774-490-1 (1-2709) x PCT-US03-21703-13 (1-341)

QY 722 CTGCTGACAGCATCCCTTTTA 742
Db 19 LeuLeuThra1aSerLeuLeu 25

RESULT 73
US-10-387-629-106
; Sequence 106, Application US/10387629
; GENERAL INFORMATION:
; APPLICANT: Chemcom S.A.
; APPLICANT: Veithen, Alex
; TITLE OF INVENTION: Olfactory and Pheromones G-Protein coupled Receptors
; FILE REFERENCE: 9409/2192
; CURRENT APPLICATION NUMBER: US/10/387,629
; PRIOR FILING DATE: 2003-03-13
; NUMBER OF SEQ ID NOS: 254
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 106
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo Sapiens
```

```
US-10-387-629-106

Alignment Scores:
Pred. No.: 89.2      Length: 345
Score: 7.00         Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80%   Indels: 0
DB: 6               Gaps: 0

US-09-774-490-1 (1-2709) x US-10-387-629-106 (1-345)

QY 36 TTACTAGCTTGTATACATCG 16
Db 167 LeuLeuSerLeuLeuThraSer 173

RESULT 74
US-10-374-780A-124
; Sequence 124, Application US/10374780A
; GENERAL INFORMATION:
; APPLICANT: Sherman, Bradley K
; APPLICANT: Riechmann, Jose Luis
; APPLICANT: Jlang, Cai-Zhong
; APPLICANT: Heard, Jacqueline E
; APPLICANT: Haake, Volker
; APPLICANT: Creelman, Robert A
; APPLICANT: Ratcliffe, Oliver
; APPLICANT: Adam, Luc J
; APPLICANT: Reuber, T. Lynne
; APPLICANT: Keddie, James
; APPLICANT: Brown, Pierre E
; APPLICANT: Pilgrim, Marsha L
; APPLICANT: Dubell III, Arnold T
; APPLICANT: Pineda, Omaira
; APPLICANT: Yu, Guo-Liang
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
; FILE REFERENCE: MEI-0047 CIP
; CURRENT APPLICATION NUMBER: US/10/374,780A
; PRIOR FILING DATE: 2003-02-25
; PRIOR APPLICATION NUMBER: 09/837,944
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/310,847
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 09/934,455
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/336,049
; PRIOR FILING DATE: 2001-11-19
; PRIOR APPLICATION NUMBER: 60/338,692
; PRIOR FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: 10/171,468
; PRIOR FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 10/225,066
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,067
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/225,068
; PRIOR FILING DATE: 2002-08-09
; NUMBER OF SEQ ID NOS: 2906
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 124
; LENGTH: 351
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: G597 (conserved domain in AA coordinates: 97-104,137-144)
US-10-374-780A-124

Alignment Scores:
Pred. No.: 89.1      Length: 351
Score: 7.00         Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78%   Indels: 0
```

DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-374-780A-124 (1-351)

QY 82 TTGAGCTCGCGCTGTCTCCCA 102

DB 295 LeuSerSerProValLeuPro 301

RESULT 75

US-10-341-434-178

; Sequence 178, Application US/10341434

; GENERAL INFORMATION:

; APPLICANT: Origene Technologies

; TITLE OF INVENTION: Differentially Regulated Prostate Cancer Genes

; FILE REFERENCE: 9U 204 205 R1

; CURRENT APPLICATION NUMBER: US/10/341,434

; CURRENT FILING DATE: 2003-07-18

; PRIOR APPLICATION NUMBER: US 60/348,164

; PRIOR FILING DATE: 2002-01-15

; PRIOR APPLICATION NUMBER: US 60/348,119

; PRIOR FILING DATE: 2002-01-15

; NUMBER OF SEQ ID NOS: 238

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 178

; LENGTH: 352

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-341-434-178

Alignment Scores:

Pred. No.: 89.1 Length: 352

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 0.80% Indels: 0

DB: 6 Gaps: 0

US-09-774-490-1 (1-2709) x US-10-341-434-178 (1-352)

QY 672 TCCAGCTTAAATATTTGCC 652

DB 324 SerSerLeuLysIleLeuSer 330

Search completed: August 3, 2003, 10:56:12
Job time : 72 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_nbp model

Run on: August 3, 2003, 10:32:29 ; Search time 24.5 Seconds
(without alignments)
9356.744 Million cell updates/sec

Title: US-09-774-490-1
Perfect score: 895
Sequence: 1 aactttatcttcatcgatg.....agcctttcttccataacc 2709

Scoring table:
OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 328717 seqs, 42310858 residues

Word size: 1

Total number of hits satisfying chosen parameters: 557304

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODE=frame+nbp.model -DEV=xjh
-Q/cgcn2_1/USPRO.spool/US09774490/runat.03082003.102928.4485/app.query.fasta.1.2887
-DB=Issued_Patents_AA -OPM=faetan -SUFFIX=oligo.ra1 -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=100 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=75 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US09774490 @CGN 1.1 30 @runat.03082003.102928.4485 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6
-Fgapext=7 -Ygapop=60 -Ygapext=60 -DELop=6 -DELext=7

Database : Issued Patents AA.*

1: /cgcn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgcn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgcn2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgcn2_6/ptodata/2/iaa/6B.COMB.pep.*
5: /cgcn2_6/ptodata/2/iaa/PTUS.COMB.pep.*
6: /cgcn2_6/ptodata/2/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	771	86.1	771	1	US-08-121-713D-54
2	771	86.1	771	1	US-08-835-268-54
3	771	86.1	771	2	US-09-060-692-54
4	771	86.1	771	3	US-08-833-391-54
5	771	86.1	771	4	US-09-060-610-54
6	771	86.1	771	5	PCT-US94-10151A-54
7	655	73.2	655	4	US-08-556-422A-3
8	434	48.5	477	1	US-08-136-922-2
9	13	1.5	775	4	US-09-308-179B-1
10	10	1.1	606	3	US-09-041-236-4
11	10	1.1	606	4	US-09-771-467C-4
12	9	1.0	112	4	US-09-341-461-34

13	9	1.0	295	4	US-08-556-422A-6	Sequence 6, Appl1
14	9	1.0	425	4	US-08-556-422A-7	Sequence 7, Appl1
15	9	1.0	607	4	US-08-556-422A-4	Sequence 4, Appl1
16	9	1.0	630	4	US-09-252-991A-32350	Sequence 32350, A
17	9	1.0	862	4	US-08-556-422A-2	Sequence 2, Appl1
18	9	1.0	913	4	US-08-971-089-4	Sequence 4, Appl1
19	8	0.9	57	4	US-09-152-060-58	Sequence 58, Appl1
20	8	0.9	139	4	US-09-252-991A-24681	Sequence 24681, A
21	8	0.9	332	4	US-09-252-991A-32158	Sequence 32158, A
22	8	0.9	360	4	US-09-252-991A-20193	Sequence 20193, A
23	8	0.9	374	1	US-08-464-148-2	Sequence 2, Appl1
24	8	0.9	374	1	US-08-385-500-2	Sequence 2, Appl1
25	8	0.9	374	1	US-08-846-784-2	Sequence 2, Appl1
26	8	0.9	454	3	US-09-080-044-4	Sequence 4, Appl1
27	8	0.9	454	4	US-09-531-857A-4	Sequence 4, Appl1
28	8	0.9	673	4	US-09-252-991A-25962	Sequence 25962, A
29	8	0.9	849	4	US-09-157-257-4	Sequence 4, Appl1
30	8	0.8	17	1	US-08-644-456-1	Sequence 1, Appl1
31	8	0.8	28	1	US-08-164-151-22	Sequence 22, Appl1
32	8	0.8	37	4	US-09-574-377-27	Sequence 27, Appl1
33	8	0.8	43	4	US-09-205-258-296	Sequence 296, App
34	8	0.8	46	4	US-09-369-247-82	Sequence 82, Appl1
35	8	0.8	48	4	US-09-489-847-175	Sequence 175, App
36	8	0.8	62	4	US-09-328-352-7261	Sequence 7261, Ap
37	8	0.8	62	6	5320958-24	Patent No. 5320958
38	8	0.8	67	4	US-09-107-532A-4497	Sequence 4497, Ap
39	8	0.8	67	4	US-08-456-647B-31	Sequence 31, Appl1
40	8	0.8	68	2	US-08-237-401A-31	Sequence 31, Appl1
41	8	0.8	63	4	US-09-134-001C-5401	Sequence 5401, Ap
42	8	0.8	73	4	US-09-107-532A-5492	Sequence 5492, Ap
43	8	0.8	77	2	US-08-469-537A-47	Sequence 47, Appl1
44	8	0.8	77	2	US-08-469-537A-48	Sequence 48, Appl1
45	8	0.8	81	2	US-09-014-869-32	Sequence 32, Appl1
46	8	0.8	88	4	US-09-732-210-270	Sequence 270, App
47	8	0.8	90	4	US-09-252-991A-19013	Sequence 19013, A
48	8	0.8	92	4	US-09-370-838-211	Sequence 211, App
49	8	0.8	98	4	US-09-252-991A-28763	Sequence 28763, A
50	8	0.8	101	4	US-09-107-532A-5792	Sequence 5792, Ap
51	8	0.8	103	3	US-08-905-223-437	Sequence 437, App
52	8	0.8	111	4	US-09-107-532A-4380	Sequence 4380, App
53	8	0.8	119	4	US-09-107-532A-3724	Sequence 3724, Ap
54	8	0.8	121	4	US-09-205-258-268	Sequence 268, App
55	8	0.8	131	4	US-09-107-532A-5741	Sequence 5741, Ap
56	8	0.8	137	4	US-09-252-991A-25510	Sequence 25510, A
57	8	0.8	139	3	US-09-120-663-4	Sequence 4, Appl1
58	8	0.8	133	6	516835-18	Patent No. 516835
59	8	0.8	188	4	US-09-252-991A-18336	Sequence 18336, A
60	8	0.8	150	4	US-09-252-991A-31728	Sequence 31728, A
61	8	0.8	158	3	US-09-053-197A-19	Sequence 19, Appl1
62	8	0.8	158	4	US-09-085-761A-19	Sequence 19, Appl1
63	8	0.8	159	4	US-09-134-001C-5575	Sequence 5575, Ap
64	8	0.8	161	4	US-09-252-991A-31464	Sequence 31464, A
65	8	0.8	162	4	US-09-252-991A-20901	Sequence 20901, A
66	8	0.8	165	4	US-08-311-731A-279	Sequence 279, App
67	8	0.8	166	4	US-09-252-991A-19472	Sequence 19472, A
68	8	0.8	172	2	US-08-469-537A-86	Sequence 86, Appl1
69	8	0.8	172	4	US-09-252-991A-18791	Sequence 18791, A
70	8	0.8	177	4	US-09-252-991A-24167	Sequence 24167, A
71	8	0.8	184	3	US-09-335-409-14	Sequence 14, Appl1
72	8	0.8	184	4	US-09-568-102-14	Sequence 14, Appl1
73	8	0.8	184	4	US-09-567-969-14	Sequence 14, Appl1
74	8	0.8	184	4	US-09-568-480-14	Sequence 14, Appl1
75	8	0.8	184	4	US-09-568-486-14	Sequence 14, Appl1
76	8	0.8	184	4	US-09-568-472-14	Sequence 14, Appl1
77	8	0.8	184	4	US-09-567-899-14	Sequence 14, Appl1
78	8	0.8	184	4	US-09-328-352-6771	Sequence 6771, Ap
79	8	0.8	191	4	US-09-252-991A-18626	Sequence 18626, A
80	8	0.8	192	4	US-09-199-637A-9	Sequence 9, Appl1
81	8	0.8	198	4	US-09-252-991A-26562	Sequence 26562, A
82	8	0.8	198	4	US-09-252-991A-27711	Sequence 27711, A
83	8	0.8	200	4	US-09-252-991A-18742	Sequence 18742, A
84	8	0.8	208	4	US-09-252-991A-31710	Sequence 31710, A
85	8	0.8	214	4	US-09-252-991A-27795	Sequence 27795, A

86 7 0.8 216 4 US-09-252-991A-22665
 c 87 7 0.8 218 4 US-09-198-452A-621
 88 7 0.8 230 4 US-09-252-991A-30553
 c 89 7 0.8 232 4 US-09-134-001C-5044
 90 7 0.8 234 4 US-09-364-230-10
 91 7 0.8 234 4 US-09-199-637A-61
 c 92 7 0.8 234 4 US-09-107-532A-4817
 93 7 0.8 236 4 US-09-252-991A-26384
 c 94 7 0.8 238 3 US-09-111-470-8
 95 7 0.8 248 6 5169835-15
 96 7 0.8 258 3 US-09-035-706-3
 97 7 0.8 258 3 US-08-955-841-3
 98 7 0.8 258 4 US-09-390-425-3
 99 7 0.8 258 4 US-09-566-906-3
 c 100 7 0.8 261 4 US-09-252-991A-17370

Sequence 22665, A
 Sequence 621, App
 Sequence 30553, A
 Sequence 5044, Ap
 Sequence 10, Appl
 Sequence 61, Appl
 Sequence 4817, Ap
 Sequence 26384, A
 Sequence 8, Appl
 Patent No. 5169835
 Sequence 3, Appl
 Sequence 3, Appl
 Sequence 3, Appl
 Sequence 17370, A

ALIGNMENTS

RESULT 1

US-08-121-713D-54

; Sequence 54, Application US/08121713D
 ; Patent No. 5639856

; GENERAL INFORMATION:

; APPLICANT: Goodman, Corey S.
 ; APPLICANT: Kolodkin, Alex L.
 ; APPLICANT: Matthes, David
 ; APPLICANT: Bentley, David R.
 ; APPLICANT: O'Connor, Timothy
 ; TITLE OF INVENTION: The Semaphorin Gene Family
 ; NUMBER OF SEQUENCES: 100
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
 ; STREET: 268 Bush Street, Suite 3200
 ; CITY: San Francisco
 ; STATE: CA
 ; COUNTRY: USA
 ; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/121,713D
 ; FILING DATE: 13-SEP-1993
 ; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Osman, Richard A.
 ; REGISTRATION NUMBER: 36,627
 ; REFERENCE/DOCKET NUMBER: B94-002-1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415)343-4341
 ; TELEFAX: (415) 343-4342

; TELETYPE:

; INFORMATION FOR SEQ ID NO: 54:

; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 771 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-121-713D-54

; Alignment Scores:

; Pred. No.: 0 Length: 771
 ; Score: 771.00 Matches: 771
 ; Percent Similarity: 100.00% Conservative: 0
 ; Best local Similarity: 100.00% Mismatches: 0
 ; Query Match: 86.15% Indels: 0
 ; DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-121-713D-54 (1-771)

QY	200	ATGGCGCTGTTAACTAGATTTCTCTTTCTTGGAGATTAATTAAGACAGACGA	259
DB	1	MecilylrhpneutrraragllvalCysleuphetrrpilyalleuutrrrllaargala	20
QY	260	AACATCAAGAAATGGGAAGAACATATGTCACAAAGCTGAATTTATCTTCAAGAAATGTTG	319
DB	21	AsnlyrghnnglylysaasnavaLProargheulyseueryrlyvgluemelcu	40
QY	320	GAATCCAAATGATGATCACTTTCAATGGCTGGCCCAAGCTCAGTTCATATACCTTC	379
DB	41	Gluersnbnvalletrrpneannglyleuulaaanserterrrhsthrphe	60
QY	380	CTTTGGATGAGAACGAGTAGGCTGTATGTTGACAAAGATCAATATTTTCAATTC	439
DB	61	LeuenuapglugluargserrargleutyValglylatalysaphstleupeserPhe	80
QY	440	GACCTGGTAAATATCAAGATTTTCAAAAGATTGTGGCCAGTATCTTACACAGAGA	499
DB	81	AspleuvalasnilleuysaphegnllyleValtrProvalserrrythrargarg	100
QY	500	GATGAATGCAAGTGGGCTGGAAGAACATCTGTAAGATGTGCTAATTTTCATCAAGTA	559
DB	101	Aspelycyslystrpalaaglylysaasplleuulyselysalaaasnpheilleysal	120
QY	560	CTTAAAGCATATATCAAGTCACTTGTAGCGCTGTGAACGGGGCTTTTCAATTC	619
DB	121	LeuylsalatyransnglntrrhstleutyAlaCysellythrghlyalaphensthrile	140
QY	620	TGCACCTACATTTGAATTTGACATCTCTGAGACAAATTTTAAAGCTGAGAACTCA	679
DB	141	Cythrlyrrilegullieglyhsthsprcgluaaspsanillepelylseugluanser	160
QY	680	CATTTTAAAGCGCGCTGGGAAGAGTCATATGACCTTAAGCTGACAGACATCCCTT	739
DB	161	HspleghlunnglyargglylysserProlyrasProlyrsleuutrrrllaaserleu	180
QY	740	TTAATAGATGAGATATATCTCTGAACTGCAAGCTGATTTTATGGGCGCAGACTTGT	799
DB	181	Leuileaspglyluleutyrrserglythrallaalaasphetctglyaaspphala	200
QY	800	ATCTTCGAACTCTTGGGCAACACACCCATATGACAGACAGACAGATGATTCAGGTG	859
DB	201	Illepharghrrleuglyhsthsstprcillargthrghluginhlsaspserrargtrp	220
QY	860	CTCAATATCCAAAGTTCATTAAGGCCACCTCATCTCAGAGAGTGAACAATCCGAAGAT	919
DB	221	LeuasnaspProlyspheilleseralahaaleuilesergluerraspsanProgluasp	240
QY	920	GACAAAGTATATCTTTCTTCGTAATGCAATGATGAGAAACACTGTGAAAAAGCT	979
DB	241	Aspylevaltyrrphephephearggluaasnlalleaspglylwhstserglylyala	260
QY	980	ACTGACGCTGATATGATCGATGATGCAAGAAATGACTTTTGGAGGGCAGAAATCTGGTG	1039
DB	261	ThrrhsllaargllieglylnilleCylsaasnaspsheglylyhlsasgserrleval	280
QY	1040	AATAATGGAACATCTCTCAAGCTCGCTGATTTGCTCAGTGGCCAGGTCCAAAATGGC	1099
DB	281	AsnlystrprrhrtrhpheuleuylsalargleuulieCyserrvalProglyProasngly	300
QY	1100	ATTGACACTGATTTTGAAGTCAAGTCAAGATGTAATCTTAATGAATTTAAAGATCTCTAA	1159
DB	301	IlleasphrrhsrphasapgluleuglnaspsvalPheleuutrrasnphelysaaspprolys	320
QY	1160	AATCCAGTTGTAATGAGAGTGTTTAGACTTTCAGTAACATTTTCAAGGATCAAGCCGTG	1219
DB	321	Asprrrovalatlyrrglyvalpneutrrhstserasnllephelysglyseralalaal	340
QY	1220	TGTATGATATGATGATGATGATGAGAGGGGTCTCTGTCATATGCCCACAGGGAT	1279
DB	341	CyswelyrrisermetseraspsvalArgargvalPheleuuliyProlyrralahaasargasp	360
QY	1280	GGACCCCAATATCAATGGTGGCTTATCAAGAAAGATCCCTATCAACGGCCAGAACT	1339

Db 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
Qy 320 GAATCAACAAGTGGATCACTTGAATGGCTGGCCAAAGCTCCAGTTATCATACCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60
Qy 380 CTTTGGATGAGAACGAGTAGGCTGTATGTTGAGCAAAAGATCATATTTTTCATTC 439
Db 61 LeuLeuAspGluGlnArgSerArgLeuTyrValGlyAlaLysAspHisIlePheSerPhe 80
Qy 440 GACCTGGTTATATCAAGAGATTTTCAAAAGATGTGTGGCCAGTATCTTACACCAAGA 499
Db 81 AspLeuValAsnIleLysAspPheGlnLysIleValTyrProValSerTyrThrArgArg 100
Qy 500 GATAAAGCAAGTGGGCTGAAAAAGACATCCTGAAGAAGATGGTAAATTTTATCAAGTA 559
Db 101 AspGluCysLysTrpAlaGlyLysAspIleLeuLysGluCysValaAsnPheIleLysVal 120
Qy 560 CTTAAGGCATTAATCAAGACTCACTGTACGCTGTGGAAAGGGGGCTTTTCATCCAAAT 619
Db 121 LeuLysAlaTyrAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
Qy 620 TGCACTCACTATGTAATGAGATCATCTCGAGACAAATATTTTAACTGGAGAACTCA 679
Db 141 CysThrTyrIleGluIleGlyHisHisAspGlnAspAsnIlePheLysLeuGluAsnSer 160
Qy 680 CATTTGAAAACGGCCGTGGGAGAGTCCATATGACCTTAAGCTGAGACGATCCCTT 739
Db 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuMetHisAlaSerLeu 180
Qy 740 TTAATAGATGAGAAATTATACTCTGGAACCTGCACTGATTTTATGGGCGAGACTTGTCT 799
Db 181 LeuIleAspGlyGlnLeuTyrSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200
Qy 800 ATCTTCGGAACCTCTGGGCACACCAACCAATCGAGACAGCAGCATGATCCAGGTGG 859
Db 201 IlePheArgThrLeuGlnLysHisHisAspProIleArgThrGlnGlnHisAspSerArgTyr 220
Qy 860 CTCATATATCCAAAGTTCATAGTGGCCACTCATCTCAGAGATGCAATCCCTGAAGAT 919
Db 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGluSerAspAsnProGlnAsp 240
Qy 920 GACAAGATATCTTTTCTTCCTCGTGAATGCAATAGATGAGAACTCTGGAAGAACT 979
Db 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGlnHisSerGlyLysAla 260
Qy 980 ACTCAGCCTGAATAGGTCAATATGCAAGATGATCTTGGAGGCGACAGAACTGTGGTG 1039
Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlnHisArgSerLeuVal 280
Qy 1040 AATTAAGTGAACAATTCCTCAAGCTCGTGTGATTTGCTCAGTGGCCAGGTCGAATGGC 1099
Db 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyTrpAsnGly 300
Qy 1100 ATTGACATCATTTTGAATGAACTGCAAGATGTATTTCTTAATGAATTTAAAGATCTTAA 1159
Db 301 IleAspThrHisPheAspGlnLeuGlnAspValPheLeuMetAsnPheLysAspProLys 320
Qy 1160 AATCAGTGTATATGAGTGTTTAACAATTCCTCAAGTAACTTTTCAAGGATCAAGCTGTG 1219
Db 321 AspProValValTyrGlyValPheThrHisSerSerAsnIlePheLysGlySerAlaVal 340
Qy 1220 TGTATGATAGCATGAGATGATGAGAAAGGTGTCTGTGGTCAATAGCCCAAGGAT 1279
Db 341 CysMetLysSerMetSerAspValArgArgValPheLeuGlyProTyrAlaHisArgAsp 360
Qy 1280 GGAACCAACTATCAATGGGTGCTTATCAAGAAAGAGTCCCTTACACGGCCAGAACT 1339
Db 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
Qy 1340 TGTCCAGCAAAACATTTGGTGTGTTGACTCTACAAAGACCTTCTGATGATGTTATA 1399
Db 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400

Qy 1400 ACCTTGCAAGAGATCATCCAGCCATGTACAATCCAGTGTTCCTATGAACAATCGCCCA 1459
Db 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetAsnAsnArgPro 420
Qy 1460 ATAGTATCAAAACGAGATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1519
Db 421 IleValIleLysThrAspValaLeuTyrGlnPheThrGlnIleValValAspArgValAsp 440
Qy 1520 GCAGAAAGTGAAGATAGATGTTATGTTATGGAACAGATGGTGGACCGTCTTAA 1579
Db 441 AlaGluAspGlyGlnTyrAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460
Qy 1580 GTAGTTTCATTTCTTAAGGAGACTGTATGATTTTGAAGAGGTTTGTGTCGAAGAAATG 1639
Db 461 ValValSerIleProLysGlnThrTrpTyrAspLeuGlnGlnValLeuLeuGlnGluMet 480
Qy 1640 ACAAGTTTTTGGGAACCGACTGTATTTTACGAATGAGACTTTCACAAAGCAACA 1699
Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGlnGln 500
Qy 1700 CTATATATGCTTCAACGGCTGGGCTTCCAGCTCCCTTACACCGGTGTGATTTTAC 1759
Db 501 LeuTyrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
Qy 1760 GGGAAAGCGTGTGAGTGTGCTCGCCCGAGACCTTACTGTGCTGGATGGTGTCT 1819
Db 521 GlyLysAlaCysValaGluCysLysLeuAlaAspAspProTyrCysAlaTrpAspGlySer 540
Qy 1820 GCATGTTCTGCTATTTTCCACTGCAAAAGACGCAACAGACGACAGATATTAAGAAAT 1879
Db 541 AlaCysSerArgTyrPheProThrAlaLysArgArgThrArgGlnAspIleArgAsn 560
Qy 1880 GGAAGCCACTGACTGACTGTTCACTTACACCAATGATATACAGGCGCACAGCCCT 1939
Db 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580
Qy 1940 GAAGAGAGAAATCATCATGCTGATGAGAAATAGTACGACATTTTGGAAATGCAAGTCCGAAG 1999
Db 581 GluGlnArgGlyIleIleTyrGlyValGluAsnSerSerThrPheLeuGluCysSerProLys 600
Qy 2000 TGCAGAGACGCTGTCTATTTGGCAATTCAGAGCGCAAAATGAAGCGAAAGAAAGAG 2059
Db 601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgAsnGlnGlnArgLysGlnGln 620
Qy 2060 ATCAGAGTGAATATCATATCATCAGACAGATCAAGGCTTGTGTAAGTATCAACA 2119
Db 621 IleArgValaAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
Qy 2120 CAGAGGATTCAGAGCAATTAACCTGTGCAAGCGGTGGAACATGGTTCATACAACCTTT 2179
Db 641 GlnLysAspSerGlyAsnTyrLeuCysHisAlaValaGlnHisGlyPheIleGlnThrLeu 660
Qy 2180 CTTAAGGTAAACCTGGAAGTCAATTAACAACAAGCATTTTGGAGAACTTCTTCAATAAGAT 2239
Db 661 LeuLysValaThrLeuGlnValaIleAspThrGlnHisLeuGlnGlnLeuLeuHisLysAsp 680
Qy 2240 GATGATGAGATGAGTGGCTTAAAGCAAAAGAAATGCTCAATAGCATGACACTGSCCAAG 2299
Db 681 AspAspGlyLysAspLysSerLysThrLysGluMetSerAsnSerLeuThrProSerGlnLys 700
Qy 2300 GTCTGTACAGAGACTTCATGACATCAACCAACCCCAATCTCAACAGATGATGAG 2359
Db 701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGln 720
Qy 2360 TTCTGTGAACAAGTTTGAAGAAAGGACGAAACAAAGTGTGGAAAGCCAGGACATACC 2419
Db 721 PheCysGlnGlnValaTyrLysArgAspArgGlyGlnArgGlnArgProGlyHisThr 740
Qy 2420 CCAAGGAAACAGTAAACAATGGAAGCACTTACAAAGAAATTAAGAAAGTGAACAAGAGAG 2479
Db 741 ProGluAsnSerAsnLysTrpLysHisLeuGlnGlnAsnLysLysGlyLysArgAsnArgArg 760

QY 2480 ACCACGAAATTGAGAGGACCCAGAGTGTC 2512
DB 761 ThrtsGluPheGluArgAlaProArgSerVal 771

RESULT 3

US-09-060-692-54

Sequence 54, Application US/09060692

Patent No. 5935865

GENERAL INFORMATION:

APPLICANT: Goodman, Corey S.

APPLICANT: Kolodkin, Alex L.

APPLICANT: Matches, David R.

APPLICANT: Bentley, David R.

TITLE OF INVENTION: The Semaphorin Gene Family

NUMBER OF SEQUENCES: 100

CORRESPONDENCE ADDRESS:

ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

STREET: 268 Bush Street, Suite 3200

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/060.692

FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/121.713

FILING DATE: 13-SEP-1993

ATTORNEY/AGENT INFORMATION:

NAME: Osman, Richard A.

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: B94-002-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 343-4341

TELEFAX: (415) 343-4342

TELEX:

INFORMATION FOR SEQ ID NO: 54:

SEQUENCE CHARACTERISTICS:

LENGTH: 771 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-060-692-54

Alignment Scores:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-774-490-1 (1-2709) x US-09-060-692-54 (1-771)

QY 200 ATGGGCTGTTAACTAGATGTTGCTCTTTCTGGGAGATTAATTACTTACAGCAAGAGCA 259
DB 1 MetGclYTrpLeuThrArgIleValCysLeuPheTrpIleValLeuLeuThrAlaArgAla 20
QY 260 AACTATCAGATGGAAGAAACAATGTGCCAAGCTGAATTAATCTTCAAGAAATGTG 319
DB 21 AsnTyrGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTyrLysGluMetLeu 40
QY 320 GAATCCAAAGATGATCATCTTTAATGCGTGGCCCAAGCTTCAGTTATCATCTTC 379
DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTyrHisThrPhe 60

QY 380 CTTTGGATGAGGAGGAGTACGCTGTATGTTGGAGCAAGATCATATTTTCATTC 439
DB 61 LeuLeuAspGluIuIuArgSerArgLeuTyrValGlyAluLysAspHisIlePheSerPhe 80
QY 440 GACCTGGTTAATATCAAGATTTTCAAAAGATTGTGTGGCCAGTATCTTACACCAAGA 439
DB 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTyrThrArgArg 100
QY 500 GATGAATGCAAGTGGGCTGGGAAAGACATCTCGTGAAGAAATGTCCTAATTTTCAAGGTA 559
DB 101 AspGluCysLeuTyrPheAlaGlyLysAspIleLeuLysGluCysAlaAsnPheIleVal 120
QY 560 CTTAAGGATATATATGACATCTGTTAGCTGTGGAAGCGGGCTTTTCATCAAT 619
DB 121 LeuLysAlaLysArgAsnGlnThrHisLeuTyrAlaCysGlyThrGlyAlaPheHisProIle 140
QY 620 TGCACCTCATTTGAAATTTGACATCATCTGAGGACAAATTTTAACTGGGAACTCA 679
DB 141 CysThrTyrIleGluIleGlyHisHisProGluAspAsnIlePheLysLeuGluAsnSer 160
QY 680 CATTTGAAAGGCGGCGGGAAGAGTCCATATGACCTTAAGCTGTGACAGATCCCTT 739
DB 161 HisPheGluAsnGlyArgGlyLysSerProTyrAspProLysLeuLeuThrAlaSerLeu 180
QY 740 TTAATGATGAGAAATTAATCTCTGGAAGTGCAGCTGATTTTATGGGCGAGACTTGTGCT 739
DB 181 LeuIleAspGlyGluLeuTyrSerGlyThrAlaAlaAspPheGlyLysArgPheAla 200
QY 800 ATCTTCGGAATCTTTGGGACCAACCAACCAATGAGACAGAGCATGATTCAGGTGG 859
DB 201 IlePheArgThrLeuGlyHisHisProIleArgThrGluGlnHisAspSerArgTyr 220
QY 860 CTCATGATGCAAGATGCTAATTAAGTCCCACTCATCTGAGAGTACCAATCCTGAAGAT 919
DB 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGlySerAspAsnProGluAsp 240
QY 920 GACAAAGTATACTTTTCTCCGTGAAATGCAATGATGAGAAACACTCTGGAAGAGCT 979
DB 241 AspLysValTyrPhePhePheArgGluAsnAlaIleAspGlyGluHisSerGlyLysAla 260
QY 980 ACTCAGCTTGAATTAAGTCAAGATATGCAAGATGACTTTTGGAGGCGACAGAACTGTGGT 1039
DB 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisArgSerLeuVal 280
QY 1040 AATTAATGCAACAATTTCTCAAGGCTGCTGATTTGCTCAAGTGCAGAGTCCAAATGCG 1099
DB 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
QY 1100 ATTGACATCATTTTATGATGAACTGCAGATGATTAATCTTAATGAACTTTAAAGATCCTAA 1159
DB 301 IleAspThrHisPheAspGluLeuGlnAspValPheLeuMetCAsnPheLysAspProLys 320
QY 1160 AATCAGTTGTATATGAGAGTGTATACGACTTCCAGTAACATTTTCAAGGATCAGCGTG 1219
DB 321 AsnProValValTyrGlyAlaPheThrThrSerSerAsnIlePheLysGlySerAlaVal 340
QY 1220 TGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1279
DB 341 CysMetLysSerLysSerLysValAlaArgValPheLeuGlyProTyrAlaHisAspAsp 360
QY 1280 GACCAACATATCAAGGCTGCTTATCAAGGAGAGTCCCATTCAGCGGCGAGAGACT 1339
DB 361 GlyProAsnTyrGlnTrpValProTyrGlnGlyArgValProTyrProArgProGlyThr 380
QY 1340 TGTCCAGCAAAACATTTGGTGTGTTTGAATCTTCAAGAGACCTTCTGATGATGTTAT 1399
DB 381 CysProSerLysThrPheGlyGlyPheAspSerThrLysAspLeuProAspValIle 400
QY 1400 ACCTTTGCAAGAGATATCAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1459
DB 401 ThrPheAlaArgSerHisProAlaMetTyrAsnProValPheProMetCAsnAsnArgPro 420
QY 1460 ATAGTATCAAAACGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1519

Db 421 ILeValIleYsThrAspValAsnTrpGlnPheThrGlnIleValValAspArgValAsp 440
QY 1520 GCAGAAAGTGAACAGTATGATGTTATGTTTATCGGAACAGATGTTGGACCGTTCTTAA 1519
Db 441 AlaGuaSerGlyGlnTrpAspValMetPheIleGlyThrAspValGlyThrValLeuYs 460
QY 1580 GTAGTTCAATTCCTAAGAGAGCTGTATGATTTAGAAAGAGTCTGCTGGAAGAAATG 1639
Db 461 ValValSerIleProValSerGlnThrTrpTrpAspLeuGlnGluValLeuLeuGlnMet 480
QY 1640 AACGTTTTGGGAAACGAGTGTATTTCAACCAATGAGCTTTCACATAGCAGACAA 1699
Db 481 ThrValPheArgGlnProThrAlaIleSerAlaMetGluLeuSerThrIleGlnGln 500
QY 1700 CTATATATGTTGTCACAGCGCTGGGTTGCCAGCTCCCTTTACACCGGTGATATTAC 1759
Db 501 LeuYrIleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleYr 520
QY 1760 GGAAGAGCGTGTGAGTGTGCTGCGCCGAGACCTTACTGTGCTTGGGATGTTCT 1819
Db 521 GlyYsAlaCysAlaGluCysCysLeuAlaArgAspProYrCysAlaIleTrpAspIleYr 540
QY 1820 GCATGTTCTCGCTATTTTCCACGTGCAAGAGACGACAGACGACAAAGATATAAGAT 1879
Db 541 AlaCysSerArgYrPheProThrAlaIleArgAlaArgAlaArgGlnAspIleArgAsn 560
QY 1880 GGAAGACCACTGACTCATGTTGTCAGACTTACACCAATGATTAATCACCATGGCCACACCTT 1939
Db 561 GlyAspProLeuTrpHisCysSerAspLeuHisAspAspAsnHisAspIleYrHisSerPro 580
QY 1940 GAAGAGAGATCATCATGATGATGAGAGAAATGATGACATTTTGGATGAGAGTCCGAG 1999
Db 581 GlnGlnAlaGlyIleIleYrGlyValGlnAsnSerSerThrPheLeuGlnCysSerProYs 600
QY 2000 TCGCAGAGAGCGCTGTCTATTTGGCAATTCAGAGCGCAAAATGAAAGCGCAAAAGAG 2059
Db 601 SerGlnArgAlaLeuValIleTrpGlnPheGlnArgAlaGlnGlnGlnGlnGln 620
QY 2060 ATCAGAGTGAATGATCATATCATGACAGACAGATCAAGGCTTCTGCTACGATGCTACA 2119
Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuLeuArgSerLeuGln 640
QY 2120 CAGAGAGATTGAGCAATTACCTGTCAGTGGGAGGAGCAATGGGTTGATCAAACTCTT 2179
Db 641 GlnYsAspSerGlyAsnTrpLeuCysHisAlaValGlnHisGlyPheIleGlnThrLeu 660
QY 2180 CTTAAGGTAACTCGAAGATGATGACACAGAGCAATTGGAAGAATTCTTCATTAAGAT 2239
Db 661 LeuYsValThrLeuGlnValIleAspThrGlnHisLeuGlnGlnGlnGlnGlnGln 680
QY 2240 GATGATGAGATGGCTCTTAAGACCAAGAAATGTCCATATGACATGACCTTACGCAAG 2299
Db 681 AspAspGlyAspGlySerIleThrIleGlyLeuMetSerAsnSerMetThrProSerGlnYs 700
QY 2300 GTCTGATACAGAGCTTCATGACAGTCAATCAACCAACCCCATTCACACAGATGAGTAG 2359
Db 701 ValTrpTrpArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGln 720
QY 2360 TTCTGTGAACAAGTTTGGAAAAAGGACCGAAAAACAACCTCGCAAGAGGCCAGACATACC 2419
Db 721 PheCysGlnGlnValIleTrpLeuArgAspArgIleGlnArgGlnIleArgProGlyHisThr 740
QY 2420 CCAGGAAACAGTAACTAATGAGAGCACTTACAAAGAAATTAAGAAAGTGAAGAGAGG 2479
Db 741 ProGlyAsnSerAsnLeuIleTrpIleHisIleLeuGlnGlnGlnGlnGlnGlnGln 760
QY 2480 ACCACGAATTTGAGAGGAGCACCGAGAGGTGC 2512
Db 761 ThrHisGlnPheGlnArgAlaProArgSerVal 771

RESULT 4
US-08-833-391-54

Sequence 54, Application US/08833391
Patent No. 6013781
GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Matthes, David
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 100
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/833,391
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/121,713
FILING DATE: 13-Sep-1993.
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415)343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 771 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-833-391-54
Alignment Scores:
Pred. No.: 0 Length: 771
Score: 771.00 Matches: 771
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 86.15% Indels: 0
DB: 3 Gaps: 0
US-09-774-490-1 (1-2709) x US-08-833-391-54 (1-771)
QY 200 ATGGGCGGTAACTGATGATGCTGCTTTCGGGAGATTAATTAATCAGCAAGACA 259
Db 1 MetGlyTrpLeuThrArgIleValCysLeuPheTrpGlyValLeuLeuTrpAlaAspGln 20
QY 260 AACTATCAGAAATGGAAGAACAAATGTGCCAAGGCTGAATTAATCTTACAAAGAAATGTTG 319
Db 21 AsnTrpGlnAsnGlyYsAsnAsnValProArgLeuLeuLeuSerYrIleYsGlnMetLeu 40
QY 320 GAATTCACAATGTGATCACTTTCATAGGCTTGCCCAACAGCTCCAGTTATCATCTTC 379
Db 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerYrHisThrPhe 60
QY 380 CTTTGGATAGGAAGGAGTAGAGCTGTATGTTGAGCAAGGATCAATATTGATTC 439
Db 61 LeuLeuAspIleGlnIleArgSerArgLeuYrValGlyAlaYsAspHisIlePheSerPhe 80
QY 440 GACCTGGTAAATCAAGATTTTCAAAAGATTGTGCGCAGTATCTTACACCAAGACA 499

Db 81 AspleuVal)asnilleysasphneglnlyleValTrpProValSerTyThrArgArg 100
 QY 500 GATGAATGCAAGTGGGCTGAAAAAGACATCTGAAAGATGTGCTAAATTCATCAAGATA 559
 Db 101 AspgluCyslystrpAlaGlyLysAspIleleuylsglucysAlaasnPhelLysVal 120
 QY 560 CTTAAGCATTAATCAAGACTCACTTGTACCGCTGTGAAACGGGGCTTTTCATCAAT 619
 Db 121 LeuylsalatyraengIntHrHisLeuTyraLacCysglTyThrGlyAlaPhenHisProIle 140
 QY 620 TGACCCATCAATTGAATTAATGACATCATCTGAGACACAAATATTTTAAGCTGAGAACTCA 679
 Db 141 CysThrTyrlleGlnIleGlyHisHisProGluAspAsnIlePhelysLeuGluAsnSer 160
 QY 680 CATTTGAAAAACGGCGCTGGGAAGATCCATATGACCTTAAGCTGTGACAGCATCCCT 739
 Db 161 HishehuLunsglyArgGlyLysSerProTyraAspProlysbLeuIntHrHisLeu 180
 QY 740 TTAATAGATGAGAAATTAATCTCTGAACTGACAGCTGATTTTATGGGGCAGACTTGTCT 799
 Db 181 LeuileAspglyleuTyrseryGlyThrAlaAlaAspPhenMetGlyArgAspPhenAla 200
 QY 800 ATCTCCGAACTCTGGGGCACACACCAATCCAGACAGACAGACAGCATGATTCAGGTGG 859
 Db 201 IleheArghThreuglyHisHisAspProIleArgThrgluglnHisAspSerArgTrp 220
 QY 860 CTCATATGATCAAAAGTTCATTAGTGGCCACTCTCTCAGAGAGTGAACATCTCGAAGAT 919
 Db 221 LeuAsnAspProlysbPheIleSerAlaHisLeuIleSerGluSerAspAsnProGluAsp 240
 QY 920 GACAAATATATCTTTTCTTCCTGAAATGCAATAGATGAGAAACATCTGAAAAAGCT 979
 Db 241 AspyValIlyrPhePhePheArgGluAsnAlaIleAspglyHisHisSerGlyLysAla 260
 QY 980 ACTCAAGCTGATAGTGAATGCTGATGCAAGATTAATCTTGGAGGGCACAGAGCTGTGGT 1039
 Db 261 ThrHisAlaArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisHisAspSerLeuVal 280
 QY 1040 AATAAATGACAACATCTCTCAAAAGCTGCTGTGATTGCTCAGTGGCAAGGCTCAAAATGCG 1099
 Db 281 AsnlystrpThrThrPheLeuLysAlaArgLeuIleCysSerValProGlyProAsnGly 300
 QY 1100 ATTGACATCATTTTGAATGAACTGACAGATGTATTCCTAATGAACTTTAAAGATCCCTAA 1159
 Db 301 IleAspThrHisAspPheAspglyleuGlnAspValPheleuMetAsnPhelysAspProlys 320
 QY 1160 AATCCAGTGTATATGAGAGTGTTCACACTTCAGTAACTTTTCAAGGAGTCAAGCTGG 1219
 Db 321 AsnProValValTyrglyValIlePheThrThrSerSerAsnIlePhelysglySerAlaVal 340
 QY 1220 TGTATGATAGCATGAGTGTATGAGAAAGGTGTCTTGCTGCATATGCGCAAGGAT 1279
 Db 341 CysMetCysSerMetSerAspValArgArgValPheleuGlyProTyraAlaHisAspAsp 360
 QY 1280 GGAACCAACTATCAATGGGTGCTTATCAAGAAAGATCCCTATCCACGGCCAGAGACT 1339
 Db 361 GlyProAsnTyrgIntPValProTyrglnGlyArgValProTyraAspProGlyThr 380
 QY 1340 TGTCACCAAGAAAACATTTGGTGTGTTGACTCTCAAAAGACCTTCTGTATGATGTATA 1399
 Db 381 CysProSerTystrPheGlyGlyPheAspSerThryAspLeuProAspAspValIle 400
 QY 1400 ACCTTTGACAGAGATCATCAGACATGTACATCCAGATGTTTCTTGAACATTCGCCCA 1459
 Db 401 ThrPheAlaArgSerHisProAlaMetTyraAsnProValPheProMetAsnAsnArgPro 420
 QY 1460 ATAGTATCAAAAACGATGTAAATTTATCAATTTTACAAATTTGTCTGAGCCGAGTGGAT 1519
 Db 421 IleValIleIleTystrAspValAsnTyrglnPheMetGlnIleValValAspArgValAsp 440
 QY 1520 GCGAAGATGAGACATGATGATTTATTTATTCGGAACAGATGTTGGAGCCGTTCTTAA 1579
 Db 441 AlaGluAspGlyGlnTyraAspValMetPheIleGlyThraspValGlyThrValIleuyls 460

QY 1580 GTAGTTTCAATTCCTAAGAGACTTGATGATTTAGAAAGGTTCTGCGAAGAAATG 1639
 Db 461 ValValSerIleProlysbGlnIntHrTyraAspLeuGluValIleleuGlnGluMet 480
 QY 1640 ACAGTTTTTGGGAACCGACTGTATTTTCAGAAATGAGCTTTCCATAGACAGCAAA 1699
 Db 481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrLysGlnGln 500
 QY 1700 CTATATATGTTGTTCAACCGCTGGGGTTCGCCACTCCCTTACACCGGTGTATTTAC 1759
 Db 501 LeuTyrlleGlySerThrAlaGlyValAlaGlnLeuProLeuHisArgCysAspIleTy 520
 QY 1760 GGGAAACCGTGTCTGAGTGTCTGCGCCGAGACCTTACTGTGTGGTGAATGTTCT 1819
 Db 521 GlyLysAlaCysAlaGluCysCysLeuAlaArgAspProTyrcysAlaTrpAspGlySer 540
 QY 1820 GCATGTTCTGCTATTTTCCACTGCAAGAGACGCAAGACAGACAGATTAAGAAAT 1879
 Db 541 AlaCysSerArgTyraPheProThrAlaLysArgThraArgGlnAspIleArgAsn 560
 QY 1880 GGAAGCCCACTGACTACTGTTCAGACTTAACCATGATTAATCAACATGGCCACAGCCT 1939
 Db 561 GlyAspProLeuThrHisCysSerAspLeuHisHisAspAsnHisHisGlyHisSerPro 580
 QY 1940 GAAGAGAGATCATCTATGTTGTGAGAAATAGTAGACATTTTGGAAATGACATCCGAAG 1999
 Db 581 GluGlnArgIleIleTyrglyValGluAsnSerSerThrPheleuGluCysSerProlys 600
 QY 2000 TCGCAGAGACCGCTGTCTATTTGGCAATTCAGAGCGCAATGAGAGCGAAAAAGAG 2059
 Db 601 SerGlnArgAlaLeuValTyrrpGlnPheGlnArgArgAsnGluGlnArgGly 620
 QY 2060 ATCAGAGTGTATATCATCATCAGACAGATCAAGGCTTGTCTGAGTACTTCAAA 2119
 Db 621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
 QY 2120 CAGAGAGATTCAGAGCAATTAACCTCTGCAAGCGGTGGAACATGGTTCTAACAACCTT 2179
 Db 641 GlnLysAspSerGlyAsnTyrcysHisAlaValGlnHisGlyPheIleGlnIntHrLeu 660
 QY 2180 CTTAAGGTAAACCTGGAAGTCAATGACAGACATTTTGGAGAACTTTCTCAATAAGAT 2239
 Db 661 LeuLysValThrleuGlnValIleAspThrGlnHisleuGlnGluLeuLeuHisLysAsp 680
 QY 2240 GATGATGAGATGGCTCTTAAGACCAAAAGATGTCATTAAGCATCAACCTTCCAGAG 2299
 Db 681 AspAspGlyAspGlySerLysThrLysGluMetSerAsnSerMetThrProSerGlnLys 700
 QY 2300 GTCTGTACAGAGACTTCATGACGCTCATCAACCAACCCCAATCTCAACAGATGTAG 2359
 Db 701 ValTrpTyraArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetAspGlu 720
 QY 2360 TTCTGTGAACAAGTTTGAAGAAAGGACCGAAAAACAAGTCTGGCAAGGCGCAGACATACC 2419
 Db 721 PheCysGlnGlnValTrpLysArgAspArgLysGlnArgGlnArgProGlyHisThr 740
 QY 2420 CCAAGGACAGATPACAAATGAGAAACACTTAACAGAAATTAAGAAAGTGAAGACAGAG 2479
 Db 741 ProGlyAsnSerAsnLysTrpLysHisleuGlnGlnLysAsnLysGlyTyraGlnAspArg 760
 QY 2480 ACCCAAGATTTTGAAGGGGACCCAGAGAGTGC 2512
 Db 761 ThrHisGluPheGlnArgAlaProArgSerVal 771

RESULT 5

US-09-060-610-54

; Sequence 54, Application us/09060610

; Patent No. 6344544

; GENERAL INFORMATION:

; APPLICANT: Goodman, Corey S.

; APPLICANT: Kolodkin, Alex L.

; APPLICANT: Matthes, David

APPLICANT: Bentley, David R.
 APPLICANT: O'Connor, Timothy
 TITLE OF INVENTION: The Semaphorin Gene Family
 NUMBER OF SEQUENCES: 100
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
 STREET: 268 Bush Street, Suite 3200
 CITY: San Francisco
 STATE: CA
 COUNTRY: USA
 ZIP: 94104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/060,610
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/835,268
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Osman, Richard A.
 REGISTRATION NUMBER: 36,627
 REFERENCE/DOCKET NUMBER: B94-002-1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415)343-4341
 TELEFAX: (415) 343-4342
 TELEX:
 INFORMATION FOR SEQ ID NO: 54:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 771 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-09-060-610-54

Alignment Scores:
 Pred. No.: 0 Length: 771
 Score: 771.00 Matches: 771
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 86.15% Indels: 0
 DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-060-610-54 (1-771)

QY 200 ATGGCTGGTTAACTAGATGTCGCTTTCTGGGGAGATTACTTACAGCAAGACA 259
 DB 1 MetcyltrpneuintrahgillevalCyeleubhetrtplgylvalleuLeuThmlaIrghala 20
 QY 260 AACTATCAGATGGAGAAACAATGTGCCAAGCTGAAATTATCTTCAAGAAATGTG 319
 DB 21 AsntryrGlnsngilylYsaasnValProargLeuylsLeuSeryrYlYsglmeleu 40
 QY 320 GAATCCCAACATGTGATCATCTTTCAATGGCTGGCCCAACAGCTCAAGTATCATACCTTC 379
 DB 41 GluserbsenValIlethrPheasnGlyleuAlaansSerSerSeryrYrHsIsthPhe 60
 QY 380 CTTTGGATGAGAGAAAGAGTGGCTGTATGTGGCAAGCAAGATCATATTTTTCATTC 439
 DB 61 LeuleuapglunghuargserArgneutyValalglYalalyasphsIlepheserPhe 80
 QY 440 GACCTGTTAATATCAAGATTTTCAAGAATGTGTGGCCAGTATCTTACACCAAGACA 499
 DB 81 AspleuvalaenlelysaasphglnlylelvalItrproValserYrThrArgArg 100
 QY 500 GATGAATGCAAGTGGCTGGAAAAGACATCTGAAAAGATGTGCTAATTTTCATCAAGCTA 559
 DB 101 AspgluncylslystrpalagilylYsaasplleleuylsglucysalaasnphelIlelyVal 120

QY 560 CTTAAGCATATATACAGACTCACTTGACGCTGTGGAACGGGGCTTTTCACTCAAT 619
 DB 121 LeuylvalaIytrhnglnthrlsleuYrAlaCyegilyrthrGlyalaphenIsproIle 140
 QY 620 TGAACCTACATTGAAATTGGACATCATCTGAGACACATATTTTAACTGGAGAACTCA 679
 DB 141 CyethrYrIleGlnIleGlyIshIsIsproGlnAspAsnIlePheylsleuGlnAsnSer 160
 QY 680 CATTGGAAAACGGCGTGGGAAGAGTCCATATGACCCCTAAGCTGTCGACAGATCCCT 729
 DB 161 HispheGlnsngilyrghgilylYsserProYrAspProYlsleuLeuThmlaSerLeu 180
 QY 740 TTAATAGATGAGAAATTAATCTGGAAGTCACTGAGTCAATTTTATGGGCGAGACTTGTCT 759
 DB 181 LeuIleapglYgluLeuYrserGlyThmlaIalaspheleuGlyalgaasphneIa 200
 QY 800 ATCTTCCGAACCTTTGGGACCAACCAATCAGACAGACAGCATGATTCGAGGTG 859
 DB 201 IlephetrGthrleuGlyhIsIshIsIsproIleArgThrGlnGlnhIsaasSerArgTrp 220
 QY 860 CTCGAATGATCAAGTTCAATTAAGTCCACCTCATCTCAGAGAGTGCATCTGGAAGAT 919
 DB 221 LeuasnAspProYlsphelIeserAlaIshleuIleSerGlnserAspAsnProGlnAsp 240
 QY 920 GACAAAGTATATCTTTTCTCCGTGAAAATGCAATAGATGAGAACACTCTGAAAAAGCT 979
 DB 241 AspYlYvalYrthrPhePheArgGlnAsnAlaIleaspGlylunhIsSerGlyYlYbala 260
 QY 980 ACTCAGCTTGAATAGTCAATATGCAAGATGACTTTGGAGGCGACAGAACTGTGCTG 1039
 DB 261 ThrhIsalargIleGlylnIleCylysaasnAspPheGlylYghIshAsgsSerleuVal 280
 QY 1040 AATTAATGACACATTTCCCAAGCTCGCTGATTTGCTCAATGGCCAGGTCCAATATGC 1099
 DB 281 AsnlystrIthrThrPheleuYlYbalaArgleuIleCysSerValProGlyProAsnGly 300
 QY 1100 ATTGACACTCATTTTGTATGATGACAGTCAAGATGTATTCCTAATGAACTTTAAAGTCTTAA 1159
 DB 301 IleaspThrhisPheaspGlnleuGlnAspValPheleuLeuKasnheylsAspProYls 320
 QY 1160 AATCCAGTTGTATATGAGAGTGTTCACGACTTCCAGTAACTTTTCAAGGATCAGCCGTG 1219
 DB 321 AsnProvalYalYrGlyValPheThrThrSerSerAsnIlePheylsGlySerAlaVal 340
 QY 1220 TGTATGATATGATGATGATGATGAGAGGGTCTCTGTCATATGCCCACAGGGAT 1279
 DB 341 CysmetyrSermetSerAspValArgArgValPheleuGlyProYrAlaIshAsp 360
 QY 1280 GGACCAACTATCAATGGGTGCTTATCAAGAGAGTCCCTATCCAGCGCCAGGAAT 1339
 DB 361 GlyProAsnYrGlnItrpValProYrGlnGlyArgValProYrProArgProGlyThr 380
 QY 1340 TGTCCAGCAAAAATTTGGTGTGTTCACCTTACAAAGACCTTCTGTATGATGTATTA 1399
 DB 381 CysProSerYrthrPheGlyGlyPheAspSerThrlysaasPheProAspAspValIle 400
 QY 1400 ACCTTTGCAAGATGATCAGGCATGTAACAATCCAGTGTTCCTATGAAACATTCGCCCA 1459
 DB 401 ThrPhehlaArgserhIsproAlaMetYrAsnProValPheProMetKasnAsnAspPro 420
 QY 1460 ATAGTATCAAAAACGATGTAAATTTATCAATTTTACAAATTCGTAGACCGAGTGAT 1519
 DB 421 IleValIlelystrhAspValaenYrGlnPheThrGlnIleValIalaspArgValasp 440
 QY 1520 GCAGAAAGATGACAGTATGATGTATGTTATGGAACAGATGTGGACCGTCTTAA 1579
 DB 441 AlaGlnAspGlylunYrAspValMetPheIleGlyThrAspValalGlyThrValleuYs 460
 QY 1580 GTAGTTTCATTTCTTCAAGAGACTTGTATGATTTTGAAGAGGTTCTGTCGAGAAAGAT 1639
 DB 461 ValValasertIleProYlsGlnItrpYrAspLeuGlnGlnValleuLeuGlnGlnmet 480
 QY 1640 ACAGTTTTCGGGAACGACCTGCTATTTCAAGAAATGAGCTTTCACATAAGCAGAAACA 1699

```

Db      .481 ThrValPheArgGluProThrAlaIleSerAlaMetGluLeuSerThrIleGlnGln 500
QY      1700 CTATATATTGGTTCACAGCGCTGGGGTGTGCCAGCTCCCTTTACACCGGTGTGATATTAC 1759
Db      501 LeuTyrIleGlySerThrIleGlyValAlaGlnLeuProLeuHisArgCysAspIleTyr 520
QY      1760 GGAAGAAGCGTGGCTGAGAGTGGTCCCTGCCCGGAGACCTTACCTGGTGGTGGAGTGGTCT 1819
Db      521 GlyIleValAcYSalIleGluCysLeuHisArgAspProTyrCysValatrrpAspGlySer 540
QY      1820 GCATGTTCTCGCTATTTTCCACCTGCCAAGAAGACGCGACAAACGACAAAGATATAGAAT 1879
Db      541 AlaCysSerArgTyrPheProThrAlaIleValSerArgThrArgGlnAspIleArgAsn 560
QY      1880 GGAGACCCACATGACCTCACTGTTTCAGACTTACACCATGATATACATGCGCACAGCCTT 1939
Db      561 GlyAspProLeuThrHisCysSerAspLeuHisAspAspProTyrCysValatrrpAspGlySer 580
QY      1940 GAAGAGAAATCATCATCTATGCTGTAGAGAAATAGTAGACATTTTGGATATGACGCGCAAG 1999
Db      581 GlnGluArgIleIleTyrGlyValGlnAsnSerSerThrPheLeuGluCysSerProIlys 600
QY      2000 TCGCAGAGAGCGCTGCTATTTGGCAATTCAGAGCGCAAAATGAAAGCGCAAAAGAAG 2059
Db      601 SerGlnArgAlaLeuValTyrTrpGlnPheGlnArgAlaGlnGlnIleArgIleGln 620
QY      2060 ATCAGAGAGATGATCATATGATACAGGACAGATCAAGGCGCTTGTCTACGTAAGTCAA 2119
Db      621 IleArgValAspAspHisIleIleArgThrAspGlnGlyLeuLeuLeuArgSerLeuGln 640
QY      2120 CAGAAAGATTCAAGCAATTACCTCTGCCCATCGGTGAACATGGGTTTCATCAAACTCTT 2179
Db      641 GlnIleAspSerGlyAsnTyrLeuCysHisIleValAlaGlnIleGlyPheIleGlnThrLeu 660
QY      2180 CTTAAAGGTAACCTTCGGAAGTCAATGACACACAGCATTTTGGAAAGACTTCTCATAAAGAT 2239
Db      661 LeuIleValIleIleLeuGlnValIleIleAspThrIleHisLeuGlnIleLeuHisIleAsp 680
QY      2240 GATGATGAGATGGCTCTTAAGACCAAGAAATGTCATATAGACATGACACCTAGCCAGAG 2299
Db      681 AsnAspGlyAspGlySerIleThrIleGlnMetSerAsnSerMetThrProSerGlnIlys 700
QY      2300 GTCTGGTTCAGAGACTTCATGCAAGTCAACCAACCCCAATCTCAACAGATGATGATG 2359
Db      701 ValTrpTyrArgAspPheMetGlnLeuIleAsnHisProAsnLeuAsnThrMetLeAspGln 720
QY      2360 TTCGTGTGAACAAGTTTGGAAAAAGGACCGGAAACAAACAGTGGCGAAAGCGCAGACATACC 2419
Db      721 PheCysGlnGlnValIleTrpIleArgAspArgIleGlnArgGlnAspProGlyHisIleThr 740
QY      2420 CCAGGGAAACAGTAACAAAATGAGACACTTACAGAAATAATGAAGGTAGAAAGAGAGG 2479
Db      741 ProGlyAsnSerAsnLeuIleTrpIleHisIleLeuGlnGlnAsnLeuIleGlyArgAsnArg 760
QY      2480 ACCCAACGAATTTGAGAGGGCAACCCAGAGATGTC 2512
Db      761 ThrHisGlnPheGlnArgAlaProArgSerVal 771

RESULT 6
PCT-US94-10151A-54
Sequence 54, Application PC/TUS9410151A
GENERAL INFORMATION:
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR HOHBACH TEST ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-4187

```

```

1  COMPUTER READABLE FORM:
2  MEDIUM TYPE: Floppy disk
3  COMPUTER: IBM PC compatible
4  OPERATING SYSTEM: PC-DOS/MS-DOS
5  SOFTWARE: Patent Release #1.0, Version #1.25
6  CURRENT APPLICATION DATA:
7  APPLICATION NUMBER: PCT/US94/10151A
8  FILING DATE: 13-SEP-1994
9  CLASSIFICATION:
10 ATTORNEY/AGENT INFORMATION:
11 NAME: Osman, Richard A.
12 REGISTRATION NUMBER: 36,627
13 REFERENCE/DOCKET NUMBER: PP-58750-PC/B/AO
14 TELECOMMUNICATION INFORMATION:
15 TELEPHONE: (415) 781-1989
16 TELEFAX: (415) 398-3249
17 TELEX: 910 277299 FHT UR
18 INFORMATION FOR SEQ ID NO: 54:
19 SEQUENCE CHARACTERISTICS:
20 LENGTH: 771 amino acids
21 TYPE: amino acid
22 TOPOLOGY: linear
23 MOLECULE TYPE: protein
24 PCT-US94-10151A-54

```

Alignment Scores:

Pred. No.:	0	771
Score:	771.00	771
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	86.15%	Indels: 0
DB:	5	Gaps: 0

US-09-774-490-1 (1-2709) X PCT-US94-10151A-54 (1-771)

Qy	200	ATGGGCTGGTAACTAGAAATGTCTGTCTTTCTCGGGAGATTAACAGCAAGACA	255
Db	1	MeGcIYTrpleuthrArgIeValCysleuethrpsIyAlleuethrAlaArgAla	20
Qy	260	AACTATCAGATGGGAAGAACATGTGCCAAGCTGAATATTCCTACAAAGAAATGTG	315
Db	21	AsnTYrGlnAsnGlyAsnAsnAlaProArgleuSleuSerTYrLysGlnMetleu	40
Qy	320	GAATCCAAAGATGTATCACTTCAATGGCTTGGGCAACACTCCAGTTATCAACCTTC	375
Db	41	GlnSerAsnAsnAlaIleThrPheAsnGlyleuAlaAsnSerSerTYrHisThrPhe	60
Qy	380	CTTTGGATGAGGAACGAGTAGGCTGTATGTTGGACAAAGATCATATTTTCATTTC	435
Db	61	LeuLeuAspGlnGlnArgSerArgleuTYrAlaGlyAlaValAspHisAlaIlePheSerPhe	80
Qy	440	GACCTGTTAATATCAAGATTTTCAAAAGATTGTGTGCCAGATCTTATACCAAGAGA	495
Db	81	AspLeuValAsnAlaLeuAspPheGlnGlyAlaValAlaTrpProValSerTYrThrArgArg	100
Qy	500	GAGGAATGCAAGGGCTGAAAAGACATCCGAAAGAAATGGTAATTTTCATCAAGGTA	555
Db	101	AspGlnCysLeuTrpAlaGlyLysAspIleLeuLysGlnCysAlaAsnPheIleLysVal	120
Qy	560	CTTAAAGCATATATCAGACTCACTTGTATACGCTGTGGAAACGGGGCTTTTCATCCAAAT	615
Db	121	LeuLYrAlaIeTYrAsnGlnThrHisLeuTYrAlaCysGlnThrGlyAlaPheHisProIle	140
Qy	620	TGACCTAATATGAATTTGGAATATATCTCTGAGACAAATATTTTAAAGCTGGAGAATCA	675
Db	141	CysThrTYrIleGlnIleGlyHisIleAspGlnAspAsnIlePheLysLeuGlnAsnSer	160
Qy	680	CATTTTGAAAAGCGCCGCTGGGAAGAGTCATATACCTTAAGCTGTGACAGATCCCTT	735
Db	161	HisPheGlnAsnGlyArgGlyLysSerProTYrAspProLysLeuLeuThrAlaSerLeu	180
Qy	740	TTAATGATGAGAAATTAATCTCTGAACTGACGTGATTTTATGGGCGAGACTTTGCT	795

[illegible]

QY	1880	GGAAACCCACGTGACTCTGTTCAGACTTTCACCATATATATACACATGGCCACAGCCCT	1993
Db	561	GLYAPRProLeuThrHisCySerSerAspLeuHisHisAspSerHisHisGlyHisSerPro	580
QY	1940	GAACAGAGAAATCATCTATGTGTGTAGAAATATAGTACCATTTTGTGAATGCACTCCGAG	1999
Db	581	GIUGIunArgIleIleTyrgIyValGIunAsnSerSerThrPheLeuGIuCySerProIys	600
QY	2000	TTCGACAGAGCGCGTGTGCTTATTTGGCAATTCACAGGCGGAAATTAAGAGCCAAAAGAG	2059
Db	601	SerGIunArgIleLeuValIyTyrgGIunPheGIunArgHisnGIunArgIySGIuGIu	620
QY	2060	ATCAGAGTGGATGATCATATCATATCAGAGACAGATCAAGACCCCTTCTGTAAGTACTCAAA	2119
Db	621	ILeArgValAspAspHisIleIleIleuTyHisAspGIunIyLeuLeuLeuArgSerIleuGIu	640
QY	2120	CAGAAGGATTCAGGCAATTACCTCTGCGCATGCGGTGGAAACATGGGTTCATACAAACTCTT	2179
Db	641	GIunIyAspSerGIuAnTyrgLeuCyHisIalIalValGIunHisGIyPheIleGIunThrIleu	660
QY	2180	CTTAAAGTAAACCCCTGGAAGTCATTTGACACAGAGATTTTGGAAAGAACTTCTTCATTAAGAT	2239
Db	661	LeuIyValThrIleuGIuValIleIleAspThrGIunHisIleuGIuGIuIleuLeuHisIyAsp	680
QY	2240	GATGATGAGATGCGCTCTTAAGACCAAGAAATGCTCCATATGACATGACACCTAGCCAGAAAG	2299
Db	681	AspAspArgIyAspArgIySerIyThrIyGIuMetSerHisSerHisMetThrProSerGIunIys	700
QY	2300	GTCGTGTACAGAGACTTCATGCACTGACGCTCATCAACCAACCCCAATCTTCAACAAGATGGATGAG	2359
Db	701	ValIyTyrgIyAspArgPheMetGIunLeuIleAsnHisProAsnIleuAnThrMetAspGIu	720
QY	2360	TTCTGTGTAACAAGTTTGAAAGAGGACCGAAACCAAGTGGCGCAAGGCCACAGACATATACC	2419
Db	721	PheCySGIunGIunValIyTyrgIyAspAspArgIySGIunArgGIunArgProGIunHisIleThr	740
QY	2420	CCAGGGAACAGTAAACAATGGAGAGCACTTACAGAGAAATATAGAAAGGTAGAAACAGAGNG	2479
Db	741	ProGIunAsnSerIleuAnIySTyrgIyHisIleuGIunGIunIleuIySGIyArgAsnArgArg	760
QY	2480	ACCACGCAATTTGAGAGGGCACCCACGAGAGTCTC	2512
Db	761	ThrHisGIunPheGIunArgIalProArgSerVal	771
RESULT 7			
US-08-556-422A-3			
; Sequence 3, Application US/08556422A			
; Patent No. 6576754			
; GENERAL INFORMATION:			
; APPLICANT: HALL, Kathryn T.			
; APPLICANT: FREEMAN, Gordon J.			
; APPLICANT: SCHULTZE, Joachim L.			
; APPLICANT: BOUSSIOTIS, Vassiliki			
; APPLICANT: NADLER, Lee M.			
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES			
; FILE REFERENCE: DEN-005CPA2			
; CURRENT APPLICATION NUMBER: US/08/556,422A			
; CURRENT FILING DATE: 1995-11-09			
; NUMBER OF SEQ ID NOS: 7			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 3			
; LENGTH: 655			
; TYPE: PR1			
; ORGANISM: Homo sapiens			
US-08-556-422A-3			
Alignment Scores:			
Pred. No.: 0			
Score: 655.00			
Percent Similarity: 100.00%			
Best Local Similarity: 100.00%			
Query Match: 73.18%			
Length: 655			
Matches: 655			
Conservative: 0			
Mismatches: 0			
Indels: 0			

DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-556-422A-3 (1-655)

QY 200 ATGGCGTGTAACTAGGATGCTGCTCTTCTGGGAGATTAATCTTACAGACAGA 259

DB 1 MetcyltrpLeuThrAlaValCysLeuPheTrpGlyValLeuLeuThrAlaArgAla 20

QY 260 AACTACAGAAATGGGAGAACAAATGTCACAGGCTGAAATTAATCTTCAAGAAATGTG 319

DB 21 AsnTrpGlnAsnGlyLysAsnAsnValProArgLeuLysLeuSerTrpLysGlnMetLeu 40

QY 320 GAATCCAAACAATGTGATCACTTCAATGCGCTGGCCCAACAGCTCCAGTATCATACCTTC 379

DB 41 GluSerAsnAsnValIleThrPheAsnGlyLeuAlaAsnSerSerTrpHisThrPhe 60

QY 380 CTTTGTGATGAGAGAGAGAGCTGTATGTTGAGCAAAAGATTCATATTTTCATTC 439

DB 61 LeuLeuAspGlnGluAspSerArgLeuTrpValGlyAlaLysAspHisIlePheSerPhe 80

QY 440 GACCTGTTAATATCAAGATTTTCAAAAGATTTGTGGCCAGTATCTTACACAGAGA 499

DB 81 AspLeuValAsnIleLysAspPheGlnLysIleValTrpProValSerTrpThrArgArg 100

QY 500 GATGATGCAAGTGGCTGAGAAAAGACATCTCGAAGAAATGTGCTAAATTTTCATCAAGTA 559

DB 101 AspGlnCysLysTrpAlaGlyLysAspIleLeuLysGlnCysAlaAsnPheIleLysVal 120

QY 560 CTTAAGGCATATATATCAAGTCTTGTAGCCCTGTGGAACGGGGGCTTTTCATCCAT 619

DB 121 LeuLysAlaTrpAsnGlnThrHisLysLeuTrpAlaCysGlyThrGlyAlaPheHisProIle 140

QY 620 TGCACCAATGTAATTAATGATCATCTGAGAGACATATTTTAACTGGAGAACTCA 679

DB 141 CysThrTrpIleGlnIleGlyHisIleAspGlnAspAsnIlePheLysLeuGlnLysSer 160

QY 680 CATTTGAAAAACGCGCTGGAGAGATTCATATGACCTTAAGCTGTGACAGCATCCCT 739

DB 161 HisPheGlnAsnGlyArgGlyLysSerProTrpAspProLysLeuLeuThrAlaSerLeu 180

QY 740 TTAATGATGAGAAATTAATCTCTGGAATCGACGTATTTTATGGGCGAGACTTGTCT 799

DB 181 LeuIleAspGlyGlnLeuTrpSerGlyThrAlaAlaAspPheMetGlyArgAspPheAla 200

QY 800 ATCTCCGAATCTTGGGACACCAACCAATCAGACAGACAGACAGATGATTCAGAGTGG 859

DB 201 IlePheArgThrLeuGlnGlyHisIleSerProIleArgThrGlnGlnHisAspSerArgTrp 220

QY 860 CTCATGATCAAAAGTTGATTAGTCCCACTCATCTCAGAGAGTGAACAATCTCGAAGAT 919

DB 221 LeuAsnAspProLysPheIleSerAlaHisLeuIleSerGlnSerAspAsnProGlnAsp 240

QY 920 GACAAAGTATATCTTTTCTCCGTAAGAAATGCAATGATGAGAAACACTTGGAAAACT 979

DB 241 AspLysValLysTrpPhePheArgGlnAsnAlaIleAspGlyGlnHisSerGlyLysAla 260

QY 980 ACTCAGCCTGAGATAGATGATGAGAGATGAGATGAGATGAGAGAGAGAGAGAGAGTGG 1039

DB 261 ThrHisIleAsnArgIleGlyGlnIleCysLysAsnAspPheGlyGlyHisAspSerLeuVal 280

QY 1040 AATTAATGACAAACATTTCTCAAGACTGCTGTGATTTGTCTCAGTGCAGGCTCCAAATGCG 1099

DB 281 AsnLysTrpThrThrPheLeuLysAlaArgLeuIleLysSerValProGlyProAsnGly 300

QY 1100 ATTGACACTCATTTTGTGATGACGACGAGATGATTTCTTAATGAACCTTAAAGATCCTAA 1159

DB 301 IleAspThrHisPheAspGlnLysGlnAspValPheLeuMetAsnPheLysAspProLys 320

QY 1160 AATCAGATTGATATGAGAGTGTGATGACATTCAGATTAATTTTCAAGGAGATGAGCGGTG 1219

DB 321 AsnProValValLysTrpValPheThrHisSerSerAsnIlePheLysGlySerAlaVal 340

QY 1220 TGTATGATATGATGATGATGATGAGAGAGGTGTCTTGTGCTCATATAGCCACAGGGAT 1279

DB 341 CysMetTrpSerMetSerAspValArgArgValPheLeuGlyProTrpAlaHisArgAsp 360

QY 1280 GGACCCAATCATTAATGGGGCTTATCAAGAAAGATCCCTCATCCAGGCCAGGAACT 1339

DB 361 GlyProAsnTrpGlnTrpValProTrpGlnGlyArgAlaProTrpProArgProGlyThr 380

QY 1340 TGTCCAGCAAAACATTTGGTGTGTTTGAATCTCAACAAAGAGACCTTCTGATGATGATA 1399

DB 381 CysProSerTrpThrPheGlyGlyPheAspSerThrLysAspLeuProAspAspValIle 400

QY 1400 ACCTTGCAAGAAATGATCAATCCAGCATGATACATCCAGTGTTCCTATGAACATCGCCA 1459

DB 401 ThrPheAlaArgSerHisProAlaMetTrpAsnProValPheProMetAsnAsnArgPro 420

QY 1460 ATAGTATCAAAACGATGTAATTAATCAATTTTCAACAAATTTGCTGATGACCGAGTGGAT 1519

DB 421 IleValIleLysThrAspValAsnTrpGlnPheThrGlnIleValAlaAspArgValAsp 440

QY 1520 GCAGAAAGATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1579

DB 441 AlaGlnAspGlyGlnTrpAspValMetPheIleGlyThrAspValGlyThrValLeuLys 460

QY 1580 GATGTTCAATTCCTAAGAGACTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1639

DB 461 ValValSerIleProLysGlnTrpTrpArgAspLeuGlnValLeuLeuGlnGlnMet 480

QY 1640 ACAATTTTGGGAGAACCGCATGATTTTCAAGCAATGAGATTTTCAAGCAAGCAAGCA 1699

DB 481 ThrValPheArgGlnProThrAlaIleSerAlaMetGlnLeuSerThrLysGlnGlnGln 500

QY 1700 CTATATATGATGATCAACGGCTGGGCTTGCACAGCTCCCTTACACGGGTGATTTTAC 1759

DB 501 LeuTrpIleLysSerThrAlaGlyAlaAlaGlnLeuProLeuHisArgCysAspIleTrp 520

QY 1760 GGGAAAGCGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1819

DB 521 GlyLysAlaCysAlaGlnCysCysLeuAlaArgAspProTrpCysAlaTrpAspGlySer 540

QY 1820 GCATGTTTCCGCTATTTTCCCATGCAAGAGACGCAACAGCAAGCAAGCAATTAAGAAT 1879

DB 541 AlaCysSerArgTrpPheProThrAlaLysArgArgThrArgGlnAspIleArgAsn 560

QY 1880 GGAGACCACTGATCAGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1939

DB 561 GlyAspProLeuThrHisCysSerAspLeuHisIleAspAsnHisIleGlyHisSerPro 580

QY 1940 GAAGAGAGATCATCTATGCTGTAGAGAAATGATGACAAATTTTGGATGACAGTCCGAAG 1999

DB 581 GlnGlnArgIleIleTrpGlyValGlnAsnSerSerThrPheLeuGlnCysSerProLys 600

QY 2000 TCGCAGAGACCGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2059

DB 601 SerGlnArgAlaLeuValLysTrpGlnPheGlnArgArgArgGlnGlnArgLysGlnGln 620

QY 2060 ATCAGAGTGAATGATATCATCATGAGACAGATCAAGGCTTCTGTAGTGTCTAACA 2119

DB 621 IleArgValAspAspHisIleIleArgTrpAspGlnGlyLeuLeuLeuLysSerLeuGln 640

QY 2120 CAGAAAGATTCAGGCAATTAATCTTGCATGCGAGGTGGAACATGAG 2164

DB 641 GlnLysAspSerGlyAsnTrpLeuGlyHisIleAlaValGlnHisGly 655

RESULT 8

US-08-136-922-2

Sequence 2, Application US/08136922

Patent No. 5416197

GENERAL INFORMATION:

APPLICANT: Raper, Jonathan A.

APPLICANT: Luo, Yuling

TITLE OF INVENTION: Compositions Which Regulate Neural

TITLE OF INVENTION: Regeneration and Methods of Making and Using the Same

NUMBER OF SEQUENCES: 2

SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 775
TYPE: PRT
ORGANISM: Mus sp:
US-09-308-179B-1

Alignment Scores:
Pred. No.: 0.000582 Length: 775
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.45% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-308-179B-1 (1-775)

QY 1778 TGTTCCTGCGCCGAGACCTTACTGTGCTTGAGATGCT 1816
DB 529 CysCysLeuAlaArgAspProTyrCySAIatrpaspGly 541

RESULT 10

US-09-041-236-4
Sequence 4, Application US/09041236
Patent No. 6225285
GENERAL INFORMATION:
APPLICANT: Luo, Yuling
TITLE OF INVENTION: Semaphorin K1
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,236
FILING DATE: March 11, 1998
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: EXEL98-001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 606 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-041-236-4

Alignment Scores:
Pred. No.: 0.466 Length: 606
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 3 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-041-236-4 (1-606)

QY 1043 AAATGACAACATTCTCAAGCTGCTGTG 1072
|||||

DB 255 LysItrpThrThrPheLeuLysAlaArgLeu 264

RESULT 11

US-09-771-467C-4
Sequence 4, Application US/09771467C
Patent No. 6583277
GENERAL INFORMATION:
APPLICANT: Luo, Yuling
Xiomel, Xu
TITLE OF INVENTION: Semaphorin K1 Polypeptides
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/771,467C
FILING DATE: 26-Jan-2001
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: EXEL98-001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 606 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:
US-09-771-467C-4

Alignment Scores:
Pred. No.: 0.466 Length: 606
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.12% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-771-467C-4 (1-606)

QY 1043 AAATGACAACATTCTCAAGCTGCTGTG 1072
DB 255 LysItrpThrThrPheLeuLysAlaArgLeu 264
|||||

RESULT 12

US-09-341-461-34
Sequence 34, Application US/09341461
Patent No. 6586389
GENERAL INFORMATION:
APPLICANT: Hammond, Timothy G.
TITLE OF INVENTION: Cudilin Protein, DNA Sequences Encoding Cudilin
FILE REFERENCE: D6148
CURRENT APPLICATION NUMBER: US/09/341,461
CURRENT FILING DATE: 2000-07-20
PRIOR APPLICATION NUMBER: PCT/US99/01259
PRIOR FILING DATE: 1999-01-21
NUMBER OF SEQ ID NOS: 40
SEQ ID NO 34

```
; LENGTH: 112
; TYPE: PRT
; ORGANISM: human
; FEATURE:
; OTHER INFORMATION: amino acid sequence of C1s CUB1 domain
US-09-341-461-34
```

```
Alignment Scores:
Pred. No.: 5.27 Length: 112
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.03% Indels: 0
DB: 4 Gaps: 0
```

US-09-774-490-1 (1-2709) x US-09-341-461-34 (1-112)

```
QY 151 TTGTGCGCCGAGAGAGTTCAACAT 125
Db 67 LeuCYeglglnArgSerSerAsnAsn 75
```

RESULT 13

```
US-08-556-422A-6
; Sequence 6, Application US/08556422A
; Patent No. 6576754
; GENERAL INFORMATION:
; APPLICANT: HALL, Kathryn T.
; APPLICANT: FREEMAN, Gordon J.
; APPLICANT: SCHULTZE, Joachim L.
; APPLICANT: BOUSSIOTIS, Vassiliki
; APPLICANT: NADLER, Lee M.
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES
; FILE REFERENCE: DEN-005CPA2
; CURRENT APPLICATION NUMBER: US/08/556,422A
; CURRENT FILING DATE: 1995-11-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 295
; TYPE: PRT
; ORGANISM: Gallus gallus
US-08-556-422A-6
```

```
Alignment Scores:
Pred. No.: 4.69 Length: 295
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 4 Gaps: 0
```

US-09-774-490-1 (1-2709) x US-08-556-422A-6 (1-295)

```
QY 1784 CTGCGCCGAGAGCCCTTACTGTGCTTG 1810
Db 287 LeuAlaArgAspProTyrCysAlaItrp 295
```

RESULT 14

```
US-08-556-422A-7
; Sequence 7, Application US/08556422A
; Patent No. 6576754
; GENERAL INFORMATION:
; APPLICANT: HALL, Kathryn T.
; APPLICANT: FREEMAN, Gordon J.
; APPLICANT: SCHULTZE, Joachim L.
; APPLICANT: BOUSSIOTIS, Vassiliki
; APPLICANT: NADLER, Lee M.
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES
; FILE REFERENCE: DEN-005CPA2
; CURRENT APPLICATION NUMBER: US/08/556,422A
; CURRENT FILING DATE: 1995-11-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 7
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-556-422A-7
```

```
Alignment Scores:
Pred. No.: 4.48 Length: 425
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 4 Gaps: 0
```

US-09-774-490-1 (1-2709) x US-08-556-422A-7 (1-425)

```
QY 1055 TTCTCAAGCTGCTGTGATTTGCTCA 1081
Db 104 PheLeuYsAlaArgLeuIleCysSer 112
```

RESULT 15

```
US-08-556-422A-4
; Sequence 4, Application US/08556422A
; Patent No. 6576754
; GENERAL INFORMATION:
; APPLICANT: HALL, Kathryn T.
; APPLICANT: FREEMAN, Gordon J.
; APPLICANT: SCHULTZE, Joachim L.
; APPLICANT: BOUSSIOTIS, Vassiliki
; APPLICANT: NADLER, Lee M.
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES
; FILE REFERENCE: DEN-005CPA2
; CURRENT APPLICATION NUMBER: US/08/556,422A
; CURRENT FILING DATE: 1995-11-09
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 607
; TYPE: PRT
; ORGANISM: Mus musculus
US-08-556-422A-4
```

```
Alignment Scores:
Pred. No.: 4.29 Length: 607
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.01% Indels: 0
DB: 4 Gaps: 0
```

US-09-774-490-1 (1-2709) x US-08-556-422A-4 (1-607)

```
QY 1784 CTGCGCCGAGAGCCCTTACTGTGCTTG 1810
Db 483 LeuAlaArgAspProTyrCysAlaItrp 491
```

RESULT 16

```
US-09-252-991A-32350
; Sequence 32350, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32350
```



```
; LENGTH: 57
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-152-060-58

Alignment Scores:
Pred. No.: 52.7      Length: 57
Score: 8.00         Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 0.89%    Indels: 0
DB: 4               Gaps: 0

US-09-774-490-1 (1-2709) x US-09-152-060-58 (1-57)

QY 716 CTTAGCTGCTGACAGCATCCCTT 739
DB 41 ProlysleuthralaSerleu 48

RESULT 20
US-09-252-991A-24681
; Sequence 24681, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24681
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24681

Alignment Scores:
Pred. No.: 47.2      Length: 139
Score: 8.00         Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 0.92%    Indels: 0
DB: 4               Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-24681 (1-139)

QY 1539 TCATCTGTCATCTTTCGATCC 1516
DB 105 SerTyrcysProSerSerAlaSer 112

RESULT 21
US-09-252-991A-32158
; Sequence 32158, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32158
; LENGTH: 322
```

```
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32158

Alignment Scores:
Pred. No.: 42.6      Length: 322
Score: 8.00         Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 0.89%    Indels: 0
DB: 4               Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-32158 (1-322)

QY 1315 AGTCCCTATCCAGCGCAGGAAC 1338
DB 242 SerProleuserThralaIaIgaSn 249

RESULT 22
US-09-252-991A-20193
; Sequence 20193, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20193
; LENGTH: 360
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20193

Alignment Scores:
Pred. No.: 42.1      Length: 360
Score: 8.00         Matches: 8
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match: 0.89%    Indels: 0
DB: 4               Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-20193 (1-360)

QY 1504 CGTAGACCGAGTGGATGCAGAGA 1527
DB 22 ArgArgProSerArgCysArgArg 29

RESULT 23
US-08-464-148-2
; Sequence 2, Application US/08464148
; Patent No. 5710026
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND
; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-3 AND CODING SEQUENCES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Stewart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,148
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/385,500
; FILING DATE: 08-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 374 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-464-148-2

Alignment Scores:
Pred. No.: 41.9      Length: 374
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.89%      Indels: 0
DB: 1      Gaps: 0

US-09-774-490-1 (1-2709) x US-08-464-148-2 (1-374)

QY 1590 TTCTAAGAGACTTGATGATT 1613
DB 282 Pheleuargargleuglymetile 289

RESULT 24
US-08-385-500-2
; Sequence 2, Application US/08385500
; Patent No. 571217
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND
; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-3 AND CODING SEQUENCES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/385,500
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 13952-21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 374 amino acids
```

```

; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-385-500-2

Alignment Scores:
Pred. No.: 41.9      Length: 374
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.89%      Indels: 0
DB: 1      Gaps: 0

US-09-774-490-1 (1-2709) x US-08-385-500-2 (1-374)

QY 1590 TTCTAAGAGACTTGATGATT 1613
DB 282 Pheleuargargleuglymetile 289

RESULT 25
US-08-846-784-2
; Sequence 2, Application US/08846784
; Patent No. 574765
; GENERAL INFORMATION:
; APPLICANT: Sprecher, Cindy A.
; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-2 AND
; TITLE OF INVENTION: CYTOPLASMIC ANTIPROTEINASE-3 AND CODING SEQUENCES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: Steuart Street Tower, One Market Plaza
; CITY: San Francisco
; STATE: California
; COUNTRY: US
; ZIP: 94105-1493
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/846,784
; FILING DATE: 30-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/385,500
; FILING DATE: 08-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Parmelee, Steven W.
; REGISTRATION NUMBER: 31,990
; REFERENCE/DOCKET NUMBER: 13952-21
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 467-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 374 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-846-784-2

Alignment Scores:
Pred. No.: 41.9      Length: 374
Score: 8.00      Matches: 8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 0.89%      Indels: 0
DB: 1      Gaps: 0

US-09-774-490-1 (1-2709) x US-08-846-784-2 (1-374)

QY 1590 TTCTAAGAGACTTGATGATT 1613
```

Db 282 Pheleuargyleuglymeille 289
RESULT 26
US-09-080-044-4
; Sequence 4, Application US/09080044
; Patent No. 6074649
; GENERAL INFORMATION:
; APPLICANT: AUDONNET, Jean-Christophe F.
; APPLICANT: BAUDU, Philippe G.
; APPLICANT: RIVIERE, Michel A.
; TITLE OF INVENTION: RECOMBINANT VACCINE CONTAINING FELINE HERPES VIRUS TYPE
; TITLE OF INVENTION: 1, PARTICULARLY FOR TREATING FELINE INFECTIOUS
; FILE REFERENCE: PERITONITIS
; CURRENT APPLICATION NUMBER: US/09/080,044
; EARLIER FILING DATE: 1998-05-15
; EARLIER FILING DATE: 1996-11-19
; EARLIER APPLICATION NUMBER: 95/14450
; EARLIER FILING DATE: 1995-11-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Feline herpesvirus 1
US-09-080-044-4
Alignment Scores:
Pred. No.: 40.9 Length: 454
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 3 Gaps: 0
US-09-774-490-1 (1-2709) x US-09-080-044-4 (1-454)
Cy 2498 CCCCTCAATTGCGGTCTCC 2475
Db 317 Proserginilearglyser 324
RESULT 27
US-09-531-857A-4
; Sequence 4, Application US/09531857A
; Patent No. 6387376
; GENERAL INFORMATION:
; APPLICANT: AUDONNET, Jean-Christophe F.
; APPLICANT: BAUDU, Philippe G.
; APPLICANT: RIVIERE, Michel A.
; TITLE OF INVENTION: RECOMBINANT VACCINE CONTAINING FELINE HERPES VIRUS TYPE
; TITLE OF INVENTION: 1, PARTICULARLY FOR TREATING FELINE INFECTIOUS
; FILE REFERENCE: 454313-2170.1
; CURRENT APPLICATION NUMBER: US/09/531,857A
; CURRENT FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 09/080,044
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: PCT/FR96/01830
; PRIOR FILING DATE: 1996-11-19
; PRIOR APPLICATION NUMBER: 95/14450
; PRIOR FILING DATE: 1995-11-30
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 454
; TYPE: PRT
; ORGANISM: Feline herpesvirus 1
US-09-531-857A-4
Alignment Scores:
Pred. No.: 40.9 Length: 454

Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 4 Gaps: 0
US-09-774-490-1 (1-2709) x US-09-531-857A-4 (1-454)
Cy 2498 CCCCTCAATTGCGGTCTCC 2475
Db 317 Proserginilearglyser 324
RESULT 28
US-09-252-991A-25962
; Sequence 25962, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25962
; LENGTH: 673
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25962
Alignment Scores:
Pred. No.: 39 Length: 673
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.92% Indels: 0
DB: 4 Gaps: 0
US-09-774-490-1 (1-2709) x US-09-252-991A-25962 (1-673)
Cy 1863 CGCTTGCGCTCTTGCAGTG 1840
Db 443 Argleuvalargleuphealaval 450
RESULT 29
US-09-157-257-4
; Sequence 4, Application US/09157257
; Patent No. 6375954
; GENERAL INFORMATION:
; APPLICANT: DUTTA, Sukanta K.
; APPLICANT: BISWAS, Biswajit
; APPLICANT: VENUGOPAL, Ramesh
; TITLE OF INVENTION: A SIZE-VARIABLE STRAIN-SPECIFIC PROTECTIVE ANTIGEN FOR
; TITLE OF INVENTION: POTOMAC HORSE FEVER
; FILE REFERENCE: 8172-9016
; CURRENT APPLICATION NUMBER: US/09/157,257
; CURRENT FILING DATE: 1998-09-18
; EARLIER APPLICATION NUMBER: 60/059,252
; EARLIER FILING DATE: 1997-09-18
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 849
; TYPE: PRT
; ORGANISM: Ehrlichia risticii
US-09-157-257-4
Alignment Scores:
Pred. No.: 37.9 Length: 849

Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.89% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-157-257-4 (1-849)

QY 515 GCTGAAAGACATCTGAAAGAA 538
Db 63 Aleglylsapleleulysglu 70

RESULT 30
US-08-644-456-1
Sequence 1, Application US/08644456
Patent No. 5747650
GENERAL INFORMATION:
APPLICANT: Kulesz-Martin, Molly F.
TITLE OF INVENTION: p53as PROTEIN AND ANTIBODY THEREFOR
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dunn & Associates
STREET: P.O. Box 96
CITY: Newtane
STATE: New York
COUNTRY: U.S.A.
ZIP: 14108
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.50 inch, 1.44 Mb storage
COMPUTER: Victor 300 SX/25 (IBM PC Compatible)
OPERATING SYSTEM: MS-DOS Version 5.0
SOFTWARE: Wordstar Professional Release 4
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/644,456
FILING DATE: 10-May-1996
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/100,496
FILING DATE: 2-Aug-1993
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Dunn, Michael L.
REGISTRATION NUMBER: 25,330
REFERENCE/DOCKET NUMBER: RPP:135C US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 433-1661
TELEFAX: (716) 433-1665
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 17
TYPE: amino acids
STRANDEDNESS: n/a
TOPOLOGY: n/a
MOLECULE TYPE: peptide
HYPOTHETICAL: no
ANTI-SENSE: no
FRAGMENT TYPE: n/a
ORIGINAL SOURCE:
ORGANISM: mouse
STRAIN: n/a
INDIVIDUAL ISOLATE: n/a
DEVELOPMENTAL STAGE: n/a
HAPLOTYPE: n/a
TISSUE TYPE: n/a
CELL TYPE: n/a
CELL LINE: n/a
ORGANELLE: n/a
IMMEDIATE SOURCE: sequenced from cDNA clone from mouse
LIBRARY: plasmid p6.3
CLONE:
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 11

MAP POSITION: p53 gene
UNITS:
FEATURE: n/a
NAME/KEY:
LOCATION:
IDENTIFICATION METHOD:
OTHER INFORMATION:
PUBLICATION INFORMATION:
AUTHORS: Kulesz-Martin et al.
TITLE: Endogenous p53 Protein Generated from Wild
TITLE: Type Alternatively Spliced P53 RNA in Mouse
JOURNAL: Mol. Cell. Biol.
VOLUME: 14
ISSUE: 3
PAGES: 1698-1708
DATE: March, 1994
AUTHORS: Han, K.A. and Kulesz-Martin, M.F.
TITLE: Alternatively Spliced p53 RNA in Transformed
TITLE: and No. 5747650mal Cells of Different Tissue Types
JOURNAL: Nucleic Acids Res.
VOLUME: 20
ISSUE: 8
PAGES: 1979-1981
DATE: 1992
AUTHORS: Arai, N. et al.
TITLE: Immunologically distinct p53 molecules generated
TITLE: by alternative splicing
JOURNAL: Mol. and Cell. Biol.
VOLUME: 6
ISSUE:
PAGES: 3232-3239
DATE: 1986
DOCUMENT NUMBER:
FILING DATE:
PUBLICATION DATE:
RELEVANT RESIDUES IN SEQ ID NO:
US-08-644-456-1

Alignment Scores:
Pred. No.: 562 Length: 17
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 1 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-644-456-1 (1-17)

QY 1302 CTTATCAGAGAGATGCCCT 1322
Db 9 leuilelysgluuserfro 15

RESULT 31
US-08-164-151-22
Sequence 22, Application US/08164151
Patent No. 5519115
GENERAL INFORMATION:
APPLICANT: Mepelli Ph.D. Claudio
APPLICANT: Derobertis, Cathy
APPLICANT: Stahl, Jerry
APPLICANT: Swerdloff Ph.D. Michael D.
APPLICANT: William Ph.D. Jon I.
APPLICANT: Everett Ph.D. Nicholas P.
APPLICANT: Bascomb Ph.D. Newell
TITLE OF INVENTION: Reverse Antimicrobial Peptides,
TITLE OF INVENTION: Antimicrobial Oligopeptides and Other Antimicrobial
TITLE OF INVENTION: Compositions, and Methods of Their Production and Use,
TITLE OF INVENTION: Particularly to Provide Protection Against Plant Pathogens
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lerner, David, Littenberg, Krumholz & Mentlik
STREET: 600 South Ave., West, Suite 300
CITY: Westfield

STATE: New Jersey
COUNTRY: United States of America
ZIP: 07090
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/164,151
FILING DATE: 09-DEC-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/649,784
FILING DATE: 01-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Teschner Esq., Michael H
REGISTRATION NUMBER: 32,862
REFERENCE/DOCKET NUMBER: Enlmount 3.0-042
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-654-5000
TELEFAX: 908-654-7866
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 28 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-164-151-22

Alignment Scores:
Pred. No.: 529 Length: 28
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-08-164-151-22 (1-28)

QY 182 TCCTTGTGCTTGTCTT 162
Db 6 SeruencyCysphaserleu 12

RESULT 32
US-09-574-377-27
Sequence 27, Application US/09574377
Patent No. 6573370
GENERAL INFORMATION:
APPLICANT: La Du, Bert
TITLE OF INVENTION: PONS And Uses Thereof
FILE REFERENCE: UM-04408
CURRENT APPLICATION NUMBER: US/09/574,377
CURRENT FILING DATE: 2000-05-19
NUMBER OF SEQ ID NOS: 34
SOFTWARE: PatentIn version 3.0
SEQ ID NO 27
LENGTH: 37
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-09-574-377-27

Alignment Scores:
Pred. No.: 511 Length: 37
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-09-574-377-27 (1-37)

QY 511 CTGCATTCATCTCTTGTCT 491
Db 12 LeuAlaPheIleSerGly 18

RESULT 33
US-09-205-258-296
Sequence 296, Application US/09205258
Patent No. 6525174
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007P1
CURRENT APPLICATION NUMBER: US/09/205,258
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06

```
;; EARLIER APPLICATION NUMBER: 60/048,974
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,883
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,897
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,898
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,962
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,963
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,877
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/048,878
;; EARLIER FILING DATE: 1997-06-06
;; EARLIER APPLICATION NUMBER: 60/070,923
;; EARLIER FILING DATE: 1997-12-18
;; EARLIER APPLICATION NUMBER: 60/092,921
;; EARLIER FILING DATE: 1998-07-15
;; EARLIER APPLICATION NUMBER: 60/094,657
;; EARLIER FILING DATE: 1998-07-30
;; NUMBER OF SEQ ID NOS: 1227
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 296
;; LENGTH: 43
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (10)
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;; NAME/KEY: SITE
;; LOCATION: (43)
;; OTHER INFORMATION: Xaa equals stop translation
US-09-205-258-296

Alignment Scores:
Pred. No.: 502 Length: 43
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-205-258-296 (1-43)

QY 2253 CCATCTCATCATCATCTTTA 2233
Db 18 ProSerProSerSerSerSerSer 24

RESULT 34
US-09-369-247-82
; Sequence 82, Application US/09369247
; Patent No. 6563992
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 44 Human Secreted Proteins
; FILE REFERENCE: P2024P1
; CURRENT APPLICATION NUMBER: US/09/369,247
; EARLIER FILING DATE: 1999-08-05
; EARLIER APPLICATION NUMBER: 60/074,118
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/074,157
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/074,137
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/074,341
; EARLIER FILING DATE: 1998-02-09
; EARLIER APPLICATION NUMBER: 60/074,141
; EARLIER FILING DATE: 1998-02-09
; NUMBER OF SEQ ID NOS: 172
```

```
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 82
;; LENGTH: 46
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: SITE
;; LOCATION: (46)
;; OTHER INFORMATION: Xaa equals stop translation
US-09-369-247-82

Alignment Scores:
Pred. No.: 497 Length: 46
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-369-247-82 (1-46)

QY 583 GTGAGTCTGATTATATGCTT 563
Db 3 ValSerLeuIleIleCysLeu 9

RESULT 35
US-09-489-847-175
; Sequence 175, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; EARLIER FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 175
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (48)
; OTHER INFORMATION: Xaa equals stop translation
US-09-489-847-175

Alignment Scores:
Pred. No.: 495 Length: 48
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-489-847-175 (1-48)

QY 2634 GAATTCAGCTGAGTTCACCC 2654
Db 30 GuileGlnLeuSerSerThr 36
```

RESULT 36
US-09-328-352-7261
; Sequence 7261, Application US/09328352
; Patent No. 6562358
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 7261
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-09-328-352-7261

Alignment Scores:
Pred. No.: 480 Length: 62
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0
DB: 4

US-09-774-490-1 (1-2709) x US-09-328-352-7261 (1-62)

QY 514 CCACTGCATTCATCTCTCT 494
Db 37 ProleuAlaphelleiser 43

RESULT 37
5320958-24
; Patent No. 5320958
; APPLICANT: INOUE, SUNIKO; HSU, MEI-YIN; EAGLE, SUSAN;
; INOUE, MASAYORI
; TITLE OF INVENTION: ISOLATED BACTERIAL REVERSE TRANSCRIPTASE
; NUMBER OF SEQUENCES: 24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/315,316
; FILING DATE: 24-FEB-1989
; SEQ ID NO: 24
; LENGTH: 62
; 5320958-24

Alignment Scores:
Pred. No.: 480 Length: 62
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0
DB: 6

US-09-774-490-1 (1-2709) x 5320958-24 (1-62)

QY 2185 CTTAAGAAGTTGTATGA 2165
Db 3 LeuylslySerLeuylTgln 9

RESULT 38
US-09-107-532A-4497
; Sequence 4497, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham

STATE: Massachusetts
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arindello, Pamela Deneka
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 4497:
SEQUENCE CHARACTERISTICS:
LENGTH: 67 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...67
SEQUENCE DESCRIPTION: SEQ ID NO: 4497:
US-09-107-532A-4497

Alignment Scores:
Pred. No.: 475 Length: 67
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0
DB: 4

US-09-774-490-1 (1-2709) x US-09-107-532A-4497 (1-67)

QY 2462 TCTTATTTCTGTAGTCT 2442
Db 15 SerTyrrheleuvalSerAla 21

RESULT 39
US-08-456-647B-31
; Sequence 31, Application US/08456647B
; Patent No. 5811516
; GENERAL INFORMATION:
; APPLICANT: Lemke Ph.D. et al., Greg E.
; TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/456,647B
FILING DATE: 02-JUN-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/237,401
FILING DATE: 02-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/884,486
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Wetherell Ph.D., John R.
REGISTRATION NUMBER: 31,678
REFERENCE/DOCKET NUMBER: 07251/007002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 68 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-456-647B-31

Alignment Scores:
Pred. No.: 474 Length: 68
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
Gaps: 0
DB: 2

US-09-774-490-1 (1-2709) x US-08-456-647B-31 (1-68)

QY 1175 GGAGTGTTCAGACTTCAGT 1195
|||||
Db 55 GlyValPheThrSerSer 61

RESULT 40
US-08-237-401A-31
Sequence 31, Application US/08237401A
Patent No. 5837448
GENERAL INFORMATION:
APPLICANT: Lemke Ph.D. et al., Greg E.
TITLE OF INVENTION: PROTEIN-TYROSINE KINASE GENES
NUMBER OF SEQUENCES: 54
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: US
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/237,401A
FILING DATE: 02-MAY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/884,486
FILING DATE: 15-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Haile Ph.D., Lisa A.
REGISTRATION NUMBER: 38,347
REFERENCE/DOCKET NUMBER: 07251/007001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 678-5070
TELEFAX: (619) 678-5099

INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 68 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-237-401A-31

Alignment Scores:
Pred. No.: 474 Length: 68
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
Gaps: 0
DB: 2

US-09-774-490-1 (1-2709) x US-08-237-401A-31 (1-68)

QY 1175 GGAGTGTTCAGACTTCAGT 1195
|||||
Db 55 GlyValPheThrSerSer 61

RESULT 41
US-09-134-001C-5401
Sequence 5401, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5401
LENGTH: 69
TYPE: PRT
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5401

Alignment Scores:
Pred. No.: 473 Length: 69
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
Gaps: 0
DB: 4

US-09-774-490-1 (1-2709) x US-09-134-001C-5401 (1-69)

QY 963 AACACTGTGAAGTACTC 983
|||||
Db 9 AsnThrLeuGluLysLeuLeu 15

RESULT 42
US-09-107-532A-5492
Sequence 5492, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts

COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneka
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5492:
SEQUENCE CHARACTERISTICS:
LENGTH: 73 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...73
SEQUENCE DESCRIPTION: SEQ ID NO: 5492:
US-09-107-532A-5492

Alignment Scores:
Pred. No.: 470 Length: 73
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0
DB:

US-09-774-490-1 (1-2709) x US-09-107-532A-5492 (1-73)
Qy 2263 GGCTTAGAGCCATCTCCATC 2243
Db 51 Glyleuargalalieserile 57

RESULT 43
US-08-469-537A-47
Sequence 47, Application US/08469537A
Patent No. 5843749
GENERAL INFORMATION:
APPLICANT: Maisondier, et al.
TITLE OF INVENTION: EHK AND ROR TYROSINE
NUMBER OF INVENTION: KINASES
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: NY
COUNTRY: U.S.A.
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/469,537A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/406,247
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: US/08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: US/07/736,559
FILING DATE: 26-JUL-1991
ATTORNEY/AGENT INFORMATION:
NAME: Kempler, Ph.D., Gail M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 47:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-469-537A-47

Alignment Scores:
Pred. No.: 467 Length: 77
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
Gaps: 0
DB:

US-09-774-490-1 (1-2709) x US-08-469-537A-47 (1-77)
Qy 1175 GGAGTGTTCAGACTTCCACT 1195
Db 57 Glyalphehrthrser 63

RESULT 44
US-08-469-537A-48
Sequence 48, Application US/08469537A
Patent No. 5843749
GENERAL INFORMATION:
APPLICANT: Maisondier, et al.
TITLE OF INVENTION: EHK AND ROR TYROSINE
NUMBER OF INVENTION: KINASES
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: NY
COUNTRY: U.S.A.
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,537A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/406,247
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: US/08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: US/07/736,559
FILING DATE: 26-JUL-1991

ATTORNEY/AGENT INFORMATION:
NAME: Kempster, Ph.D., Gail M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721
TELEX:
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 77 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-469-537A-48

Alignment Scores:
Pred. No.: 467
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.78%
DB: 2
Matches: 77
Conservative: 7
Mismatch: 0
Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-08-469-537A-48 (1-77)

QY 1175 GGAGTGTTCAGCTTCAGT 1195
Db 57 GYVALPETHRTHTSR 63

RESULT 45
US-09-014-969-32
Sequence 32, Application US/09014969
Patent No. 5965397
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Racie, Lisa A.
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYPEPTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 Cambridgepark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/014,969
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 81 amino acids
TYPE: amino acid

STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-014-969-32

Alignment Scores:
Pred. No.: 464
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.80%
DB: 2
Matches: 81
Conservative: 7
Mismatch: 0
Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-09-014-969-32 (1-81)

QY 672 TCCAGCTTAAATATTTGCC 652
Db 59 SerserleuylsileuSer 65

RESULT 46
US-09-732-210-270
Sequence 270, Application US/09732210
Patent No. 6573361
GENERAL INFORMATION:
APPLICANT: Bunkers, Greg J.
APPLICANT: Liang, Jihong
APPLICANT: Mitnick, Cindy A.
APPLICANT: Seale, Jeffrey W.
APPLICANT: Wu, Yonnie S.
TITLE OF INVENTION: Anti-fungal Proteins and Methods for Their Use
FILE REFERENCE: 38-21(15036)B
CURRENT APPLICATION NUMBER: US/09/732,210
CURRENT FILING DATE: 2000-12-07
PRIOR APPLICATION NUMBER: US 60/169,513
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: US 60/169,340
PRIOR FILING DATE: 1999-12-07
NUMBER OF SEQ ID NOS: 1753
SEQ ID NO 270
LENGTH: 88
TYPE: PRT
ORGANISM: Caenorhabditis elegans
US-09-732-210-270

Alignment Scores:
Pred. No.: 460
Score: 7.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 0.80%
DB: 4
Matches: 88
Conservative: 7
Mismatch: 0
Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-09-732-210-270 (1-88)

QY 2406 CTTGCCGACGTTGTTTCG 2386
Db 45 Leucylsargylcysphearg 51

RESULT 47
US-09-252-991A-19013
Sequence 19013, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19013
; LENGTH: 90
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19013

Alignment Scores:
Pred. No.: 458 Length: 90
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-19013 (1-90)

QY 1635 TCTTCAGACGACCTCTTCT 1615

Db 20 SerSerSerArgThrSerSer 26

RESULT 48

US-09-370-838-211
; Sequence 211, Application US/09370838
; Patent No. 6444425
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Mohamath, Roadah
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOUNDS FOR THERAPY AND DIAGNOSIS OF
; FILE REFERENCE: 210121.475C1
; CURRENT APPLICATION NUMBER: US/09/370,838
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: US 09/285,323
; EARLIER FILING DATE: 1999-04-02
; NUMBER OF SEQ ID NOS: 289
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 211
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-370-838-211

Alignment Scores:

Pred. No.: 457 Length: 92
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-370-838-211 (1-92)

QY 722 CTGCTGACGACATCCCTTTA 742

Db 19 LeuLeuThrAlaSerLeuLeu 25

RESULT 49
US-09-252-991A-28763
; Sequence 28763, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28763
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28763

Alignment Scores:
Pred. No.: 454 Length: 98
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-28763 (1-98)

QY 284 CATTTCTTCCCATTTCTGAT 264

Db 78 HicYSerSerHisSerAsp 84

RESULT 50

US-09-107-532A-5792
; Sequence 5792, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; CORRESPONDENCE ADDRESS: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Walham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneka
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781) 893-5007
; TELEFAX: (781) 893-8277
; INFORMATION FOR SEQ ID NO: 5792:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 101 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...101
; SEQUENCE DESCRIPTION: SEQ ID NO: 5792:
US-09-107-532A-5792

Alignment Scores:

Pred. No.: 452 Length: 101
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.78% Indels: 0
 DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-107-532A-5792 (1-101)

Qy 1423 CACTACATCCAGTGTTC 1443

Db 14 HieValGInserValSer 20

RESULT 51

US-08-905-223-437
 ; Sequence 437, Application US/08905223
 ; Patent No. 6222029
 ; GENERAL INFORMATION:
 ; APPLICANT: Edwards, Jean-Baptiste D.
 ; APPLICANT: Duclert, Aymeric
 ; APPLICANT: Lacroix, Bruno
 ; TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS
 ; NUMBER OF SEQUENCES: 503
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Knobbe, Martens, Olson & Bear
 ; STREET: 501 West Broadway
 ; CITY: San Diego
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 92101-3505
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy Disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: Win95
 ; SOFTWARE: Word
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/905,223
 ; FILING DATE:
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Israel, Ned A.
 ; REGISTRATION NUMBER: 29,655
 ; REFERENCE/DOCKET NUMBER:
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (619) 235-8550
 ; TELEFAX: (619) 235-0176
 ; INFORMATION FOR SEQ ID NO: 437:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 103 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: LINEAR
 ; MOLECULE TYPE: PROTEIN
 ; ORIGINAL SOURCE:
 ; ORGANISM: Homo Sapiens
 ; TISSUE TYPE: Brain
 ; FEATURE:
 ; NAME/KEY: sig_peptide
 ; LOCATION: 35..1
 ; IDENTIFICATION METHOD: Von Heljne matrix
 ; OTHER INFORMATION: score 7.4
 ; OTHER INFORMATION: seq SLLSLSLAGSG/CG
 ; US-08-905-223-437

Alignment Scores:

Pred. No.: 451 Length: 103
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.80% Indels: 0
 DB: 3 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-905-223-437 (1-103)

Qy 1248 CTTCACACTCACTGCTA 1228

Db 46 LeuleuTherSerleuMetleu 52

RESULT 52

US-09-107-532A-4380
 ; Sequence 4380, Application US/09107532A
 ; Patent No. 6583275
 ; GENERAL INFORMATION:
 ; APPLICANT: Lynn A Doucette-Stamm and David Bush
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
 ; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 7310
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 ; STREET: 100 Beaver Street
 ; CITY: Waltham
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02354
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: CD-ROM ISO9660
 ; COMPUTER: PC
 ; OPERATING SYSTEM: <Unknown>
 ; SOFTWARE: ASCII
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/107,532A
 ; FILING DATE: 30-Jun-1998
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 60/085,598
 ; FILING DATE: 14 May 1998
 ; APPLICATION NUMBER: 60/051571
 ; FILING DATE: July 2, 1997
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Arinello, Pamela Deneka
 ; REGISTRATION NUMBER: 40,489
 ; REFERENCE/DOCKET NUMBER: GTC-012
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (781)893-5007
 ; TELEFAX: (781)893-8277
 ; INFORMATION FOR SEQ ID NO: 4380:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 111 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; HYPOTHETICAL: YES
 ; ORIGINAL SOURCE:
 ; ORGANISM: Enterococcus faecium
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (8) LOCATION 1...111
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 4380:
 ; US-09-107-532A-4380

Alignment Scores:

Pred. No.: 447 Length: 111
 Score: 7.00 Matches: 7
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 0.80% Indels: 0
 DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-107-532A-4380 (1-111)

Qy 2481 GTCCTCTGTTTCTACCTTC 2461

Db 60 ValLeuLeuPheLeuProPhe 66

RESULT 53

US-09-107-532A-3724
 ; Sequence 3724, Application US/09107532A

Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NOCLETIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 3724:
SEQUENCE CHARACTERISTICS:
LENGTH: 119 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...119
SEQUENCE DESCRIPTION: SEQ ID NO: 3724:
US-09-107-532A-3724

Alignment Scores:
Pred. No.: 443 Length: 119
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0
DB: 4

US-09-774-490-1 (1-2709) x US-09-107-532A-3724 (1-119)

QY 1584 ACTACTTTAAGACGCTCCA 1564
Db 28 ThrThrleuakgrHrValPro 34

RESULT 54
US-09-205-258-268
Sequence 268, Application US/09205258
Patent No. 6525174
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007P1
CURRENT APPLICATION NUMBER: US/09/205,258
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422

EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 268
LENGTH: 131
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (131)
OTHER INFORMATION: Xaa equals stop translation
US-09-205-258-268

Alignment Scores:
Pred. No.: 438 Length: 131
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-205-258-268 (1-131)

QY 2520 GCAGCTCAGACACCTCGGT 2500
Db 122 AlalalacInthrLeugly 128

RESULT 55
US-09-107-532A-5741
Sequence 5741, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
City: Walham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneké
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 5741:
SEQUENCE CHARACTERISTICS:
LENGTH: 131 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc_feature
LOCATION: (B) LOCATION 1...131
SEQUENCE DESCRIPTION: SEQ ID NO: 5741:
US-09-107-532A-5741

Alignment Scores:
Pred. No.: 438 Length: 131
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-107-532A-5741 (1-131)

QY 30 TTAGTATCGATGCGACGTCG 50
Db 18 Leuvalillephalathrer 24

RESULT 56
US-09-252-991A-25510
Sequence 25510, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196,136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 25510
LENGTH: 137
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25510

Alignment Scores:
Pred. No.: 435 Length: 137
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-25510 (1-137)

QY 1328 CGGCGAGAACTGTCGCCAGC 1348
Db 50 ArgptroglYthrcysproser 56

RESULT 57
US-09-120-663-4
Sequence 4, Application US/09120663
Patent No. 6228644
GENERAL INFORMATION:
APPLICANT: Bogdanove, Adam J.
APPLICANT: Kim, Jihyun Francis
APPLICANT: Wei, Zhong-Min
APPLICANT: Beer, Steven V.
TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR FROM
TITLE OF INVENTION: ERWINIA AMYLOVORA, ITS USE, AND ENCODING GENE
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:

ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: P.O. Box 1051, Clinton Square
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: US/09/120,663
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/055,105
FILING DATE: 06-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/1661
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 139 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-120-663-4

Alignment Scores:
Pred. No.: 435 Length: 139
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: Gaps: 0

US-09-774-490-1 (1-2709) x US-09-120-663-4 (1-139)

QY 272 CATCTGATAGTTCCTTG 252

Db 48 HisserapSerLeuLeu 54

RESULT 58
5169835-18
PATENT NO. 5169835
APPLICANT: WAI-YEE, CHAN
TITLE OF INVENTION: PREGNANCY SPECIFIC PROTEIN APPLICATIONS
NUMBER OF SEQUENCES: 48
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/390,409
FILING DATE: 07-AUG-1989
SEQ ID NO: 18
LENGTH: 143
5169835-18

Alignment Scores:
Pred. No.: 433 Length: 143
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: Gaps: 0

US-09-774-490-1 (1-2709) x 5169835-18 (1-143)

QY 722 CTGCTGACAGCATCCCTTTTA 742

|||||

Db 19 LeuLeuThralaseLeuLeu 25

RESULT 59
US-09-252-991A-18336
Sequence 18336, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 18336
LENGTH: 148
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18336

Alignment Scores:
Pred. No.: 431 Length: 148
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-18336 (1-148)

QY 873 AGTTCATTAGTCCACCTCA 893

Db 3 SerSerLeuValProThrSer 9

RESULT 60
US-09-252-991A-31728
Sequence 31728, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 31728
LENGTH: 150
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31728

Alignment Scores:
Pred. No.: 431 Length: 150
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-31728 (1-150)

QY 2089 AGATCAAGCCTTCTGCTACG 2109

Db 92 ArgSerArgProSerAlaThr 98

|||||

RESULT 61
US-09-053-197A-19
; Sequence 19, Application US/09053197A
; Patent No. 6022952
; GENERAL INFORMATION:
; APPLICANT: Turner, Joel H.
; APPLICANT: Turner, Raymond J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PROTEIN
; TITLE OF INVENTION: SECRETION
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/053,197A
; FILING DATE: 01-Apr-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Macknight, Kamrin T.
; REGISTRATION NUMBER: 38,230
; REFERENCE/DOCKET NUMBER: UALB-03293
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-09-053-197A-19
; Alignment Scores:
; Pred. No.: 428 Length: 158
; Score: 7.00 Matches: 7
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 0.78% Indels: 0
; Gaps: 0
; DB: 3
US-09-774-490-1 (1-2709) x US-09-053-197A-19 (1-158)
QY 2207 ACAGAGCATTGGAGAGACTT 2227
Db 1 ThrluHlsleuclugluLeu 7
RESULT 62
US-09-085-761A-19
; Sequence 19, Application US/09085761A
; Patent No. 6335178
; GENERAL INFORMATION:
; APPLICANT: Weiner, Joel H.
; APPLICANT: Turner, Raymond J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR PROTEIN
; TITLE OF INVENTION: SECRETION
; NUMBER OF SEQUENCES: 77
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Medlen & Carroll, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America

ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/085,761A
; FILING DATE: 28-MAY-1998
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Peter G.
; REGISTRATION NUMBER: 32,837
; REFERENCE/DOCKET NUMBER: UALB-03356
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 158 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-09-085-761A-19
; Alignment Scores:
; Pred. No.: 428 Length: 158
; Score: 7.00 Matches: 7
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 0.78% Indels: 0
; Gaps: 0
; DB: 4
US-09-774-490-1 (1-2709) x US-09-085-761A-19 (1-158)
QY 2207 ACAGAGCATTGGAGAGACTT 2227
Db 1 ThrluHlsleuclugluLeu 7
RESULT 63
US-09-134-001C-5575
; Sequence 5575, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 5575
; LENGTH: 159
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-5575
; Alignment Scores:
; Pred. No.: 428 Length: 159
; Score: 7.00 Matches: 7
; Percent Similarity: 100.00% Conservative: 0
; Best Local Similarity: 100.00% Mismatches: 0
; Query Match: 0.80% Indels: 0
; Gaps: 0
; DB: 4
US-09-774-490-1 (1-2709) x US-09-134-001C-5575 (1-159)
QY 280 GTTCCTCCATCTGTAGT 260

Db 129 ValLeuProIleuLeuIleVal 135

RESULT 64
US-09-252-991A-31464
; Sequence 31464, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31464
; LENGTH: 161
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31464

Alignment Scores:
Pred. No.: 427 Length: 161
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-31464 (1-161)

QY 1527 TCTTTCGATCCTCGCTCT 1507
Db 56 SerSerAlaSerThrArgSer 62

RESULT 65
US-09-252-991A-20901
; Sequence 20901, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20901
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20901

Alignment Scores:
Pred. No.: 427 Length: 162
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-20901 (1-162)

QY 1635 TCTTCAGACGAACTCTTCT 1615
|||||

Db 134 SerSerArgThrSerSer 140

RESULT 66
US-08-311-731A-279
; Sequence 279, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:

; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311.731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 279:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 165 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Mycobacterium leprae
US-08-311-731A-279

Alignment Scores:
Pred. No.: 426 Length: 165
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
Gaps: 0

US-09-774-490-1 (1-2709) x US-08-311-731A-279 (1-165)

QY 2308 GTACCAAGCTTCTGCTAGG 2288
Db 143 ValProAspLeuLeuAlaArg 149

RESULT 67
US-09-252-991A-19472
; Sequence 19472, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 19472
LENGTH: 166
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19472

Alignment Scores:
Pred. No.: 425 Length: 166
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-19472 (1-166)

QY 966 ACTCTGGAAGCTACTCAG 986
Db 105 Thirdugluylsleuthr 111

RESULT 68

US-08-469-537A-86
Sequence 86, Application US/08469537A
Patent No. 5843749

GENERAL INFORMATION:

APPLICANT: Maisompierre, et al.
TITLE OF INVENTION: EHK AND FOR TYROSINE
TITLE OF INVENTION: KINASES
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill River Road
CITY: Tarrytown
STATE: NY

COUNTRY: U.S.A.
ZIP: 10591

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/469,537A
FILING DATE: 06-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: USSN 08/406,247
FILING DATE: 17-MAR-1995
APPLICATION NUMBER: USSN 08/144,992
FILING DATE: 28-OCT-1993
APPLICATION NUMBER: USSN 07/736,559
FILING DATE: 26-JUL-1991

ATTORNEY/AGENT INFORMATION:

NAME: Kempster, Ph.D., Gail M
REGISTRATION NUMBER: 32,143
REFERENCE/DOCKET NUMBER: REG 070C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7400
TELEFAX: 914-345-7721

TELEX:

INFORMATION FOR SEQ ID NO: 86:
SEQUENCE CHARACTERISTICS:

LENGTH: 172 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-469-537A-86

Alignment Scores:
Pred. No.: 423 Length: 172
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 2 Gaps: 0

US-09-774-490-1 (1-2709) x US-08-469-537A-86 (1-172)

QY 1175 GAGGTGTTTACGACTTCAGT 1195
Db 92 GlyValPheThrThrSerSer 98

RESULT 69

US-09-252-991A-18791
Sequence 18791, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

SEQ ID NO 18791

LENGTH: 172

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-18791

Alignment Scores:

Pred. No.: 423 Length: 172
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-18791 (1-172)

QY 1635 TCTTCAGCAGAACCTCTTCT 1615
Db 76 SerSerSerThrThrSerSer 82

RESULT 70

US-09-252-991A-24167
Sequence 24167, Application US/09252991A
Patent No. 6551795

GENERAL INFORMATION:

APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A

CURRENT FILING DATE: 1999-02-18

PRIOR APPLICATION NUMBER: US 60/074,788

PRIOR FILING DATE: 1998-02-18

PRIOR APPLICATION NUMBER: US 60/094,190

PRIOR FILING DATE: 1998-07-27

SEQ ID NO 24167

LENGTH: 177

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-24167

Alignment Scores:

Pred. No.: 422 Length: 177
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.80% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-252-991A-24167 (1-177)

QY 2118 TGTAGACTACGACGAGCAAG 2098
DB 153 CysArgLeuThrAlaArgAla 159

RESULT 71
US-09-335-409-14
Sequence 14, Application US/09335409
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/335,409
CURRENT FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 184
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-335-409-14

Alignment Scores:
Pred. No.: 420 Length: 184
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 3 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-335-409-14 (1-184)

QY 239 GTATTACTTACGACGAGCA 259
DB 58 ValLeuLeuThrAlaArgAla 64

RESULT 72
US-09-568-102-14
Sequence 14, Application US/09568102
Patent No. 6346404
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/568,102
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 184
TYPE: PRT
ORGANISM: Sorangium cellulosum

US-09-568-102-14

Alignment Scores:
Pred. No.: 420 Length: 184
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-568-102-14 (1-184)

QY 239 GTATTACTTACGACGAGCA 259
DB 58 ValLeuLeuThrAlaArgAla 64

RESULT 73
US-09-567-969-14
Sequence 14, Application US/09567969
Patent No. 6355457
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/567,969
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 184
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-567-969-14

Alignment Scores:
Pred. No.: 420 Length: 184
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-567-969-14 (1-184)

QY 239 GTATTACTTACGACGAGCA 259
DB 58 ValLeuLeuThrAlaArgAla 64

RESULT 74
US-09-568-480-14
Sequence 14, Application US/09568480
Patent No. 6355458
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/568,480
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 184
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-568-480-14

Alignment Scores:
Pred. No.: 420 Length: 184
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-568-480-14 (1-184)

QY 239 GTATTACTTACAGCAGAGCA 259
|||||
Db 58 ValleuLeuthrAlaArgAla 64

RESULT 75

US-09-568-486-14
Sequence 14, Application US/09568486
Patent No. 6355459
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligou, James
APPLICANT: Molnar, Istvan
APPLICANT: Zirkle, Ross
APPLICANT: Cyr, Devon
APPLICANT: Goerlach, Joern
TITLE OF INVENTION: GENES FOR THE BIOSYNTHESIS OF EPOTHILONES
FILE REFERENCE: 4-30582A
CURRENT APPLICATION NUMBER: US/09/568,486
CURRENT FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 09/335,409
PRIOR FILING DATE: 1999-06-17
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 14
LENGTH: 184
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-568-486-14

Alignment Scores:
Pred. No.: 420 Length: 184
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.78% Indels: 0
DB: 4 Gaps: 0

US-09-774-490-1 (1-2709) x US-09-568-486-14 (1-184)

QY 239 GTATTACTTACAGCAGAGCA 259
|||||
Db 58 ValleuLeuthrAlaArgAla 64

Search completed: August 3, 2003, 10:55:30
Job time: 74.5 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 17:20:41 / Search time 9580 Seconds
(without alignments)
11568.286 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709

Sequence: 1 aatctttatttaccgaltg.....agcttttttccataacc 2709

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 2888711 seqs, 2045481386 residues

Word size: 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database: GenEmbl:

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_un:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pln:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2709	100.0	2709	6 AX207154	AX207154 Sequence
2	2508	92.6	2530	6 AX743092	AX743092 Sequence
3	2508	92.6	2530	6 BD084822	BD084822 Diagnosis
4	2508	92.6	2530	6 HUMHSEM	L26081 Homo sapien
5	2508	92.6	2530	11 G31703	G31703 SMS1.973 Ex
6	2508	92.6	2601	6 AR040598	AR040598 Sequence
7	2508	92.6	2601	6 147054	147054 Sequence
8	1253	46.3	1481	6 111900	111900 Sequence 1
9	652	24.1	169408	2 AC073110	AC073110 Homo sapi
10	601	22.2	179640	9 AC006322	AC006322 Homo sapi
11	303	11.2	354	6 AX193169	AX193169 Sequence
12	220	8.1	94027	9 AC073221	AC073221 Homo sapi
13	194	7.2	108642	9 AC004451	AC004451 Homo sapi
14	160	5.9	135214	9 AC004848	AC004848 Homo sapi
15	148	5.5	169408	2 AC073110	AC073110 Homo sapi
16	117	4.3	81667	2 AC091270	AC091270 Homo sapi
17	71	2.6	2137	6 AX207159	AX207159 Sequence
18	68	2.5	2319	10 RNSIICN1	X95286 R.norvegicu
19	63	2.3	229	6 BD026302	BD026302 Sequence
20	63	2.3	110000	2 AC109862.2	Continuation (3 of
21	55	2.0	81667	2 AC091270	AC091270 Homo sapi
22	50	1.8	179667	2 AC095790	AC095790 Rattus no
23	50	1.8	273260	2 AC123888	AC123888 Rattus no
24	47	1.7	5952	10 MUSC1	D85028 Mus musculu
25	47	1.7	195335	2 AC068065	AC068065 Mus muscu
26	47	1.7	265982	2 AC121125	AC121125 Mus muscu
27	41	1.5	1998	10 MUSEMALI1	L40464 Mus musculu
28	41	1.5	2319	10 MUSC05E	L41541 Mus musculu
29	41	1.5	2913	10 MMRNASEMD	X85993 M.musculu
30	41	1.5	110000	2 AC109862.1	Continuation (2 of
31	41	1.5	184006	2 AC102452	AC102452 Mus muscu
32	41	1.5	207757	10 AC022368	AC022368 Mus muscu
33	33	1.2	195335	2 AC068065	AC068065 Mus muscu
34	33	1.2	195311	10 AC121585	AC121585 Mus muscu
35	31	1.1	4661	6 AX359936	AX359936 Sequence
36	31	1.1	4661	6 AX382150	AX382150 Sequence
37	31	1.1	6620	12 SYNMLPLN3	M28247 Retroviral
38	31	1.1	6837	12 SYNMOV2	M64754 Moloney mur
39	31	1.1	7160	6 AR302096	AR302096 Sequence
40	31	1.1	7235	6 AR302097	AR302097 Sequence
41	29	1.1	29	6 AX591134	AX591134 Sequence
42	27	1.0	3148	5 AF086761	AF086761 Danto rer
43	27	1.0	9729	6 AX107928	AX107928 Sequence
44	27	1.0	12473	6 AX107927	AX107927 Sequence
45	25	0.9	25	6 AX591129	AX591129 Sequence
46	25	0.9	1400	6 AX207158	AX207158 Sequence
47	25	0.9	244749	2 AC098444	AC098444 Rattus no
48	25	0.9	250672	2 AC096525	AC096525 Rattus no
49	24	0.9	24	6 AX591128	AX591128 Sequence
50	24	0.9	3263	5 GGU02528	UG2528 Gallus galli
51	24	0.9	225005	2 AC102985	AC102985 Rattus no
52	23	0.8	29	6 AX591132	AX591132 Sequence
53	23	0.8	1822	10 D21099	D21099 Mouse mRNA
54	23	0.8	2325	5 AY030051	AY030051 Xenopus 1
55	23	0.8	51312	2 AC100365	AC100365 Mus muscu
56	23	0.8	225372	5 AC120544	AC120544 Mus muscu
57	22	0.8	883	5 GGU028241	UG28241 Gallus galli
58	22	0.8	5732	6 AX359932	AX359932 Sequence
59	22	0.8	5732	6 AX382146	AX382146 Sequence
60	22	0.8	7311	6 AB0476	AB0476 Sequence 3
61	22	0.8	7311	6 BD136840	BD136840 Transoma
62	22	0.8	7352	6 AR302095	AR302095 Sequence
63	22	0.8	7353	6 AR302092	AR302092 Sequence
64	22	0.8	7591	6 AX107929	AX107929 Sequence
65	22	0.8	7885	6 AB0477	AB0477 Sequence 4

C	66	22	0.8	7885	6	BD136841	Transposon
C	67	22	0.8	9183	6	AX359933	AX359933 Sequence
C	68	22	0.8	9183	6	AX382147	AX382147 Sequence
C	69	22	0.8	96064	9	AC006525	AC006525 Homo sapi
C	70	22	0.8	110000	2	AC112426	AC112426 Rattus no
C	71	22	0.8	110000	2	AC112426	Continuation (3 of
C	72	22	0.8	150638	2	AC129994	AC129994 Rattus no
C	73	22	0.8	166965	2	AC122712	AC122712 Homo sapi
C	74	22	0.8	216124	2	AC098305	AC098305 Rattus no
C	75	22	0.8	223086	2	AC106106	AC106106 Rattus no
C	76	22	0.8	276158	2	AC129229	AC129229 Rattus no
C	77	22	0.8	282515	2	AC118094	AC118094 Rattus no
C	78	22	0.8	298698	2	AC094778	AC094778 Rattus no
C	79	22	0.8	328905	2	AC094646	AC094646 Rattus no
C	80	21	0.8	29	6	AX591133	AX591133 Sequence
C	81	21	0.8	618	11	BV056652	BV056652 S2126P751
C	82	21	0.8	1385	6	AX174768	AX174768 Sequence
C	83	21	0.8	1385	6	BD087170	BD087170 Method of
C	84	21	0.8	1661	6	AR217930	AR217930 Sequence
C	85	21	0.8	1661	6	BD057812	BD057812 Growth fa
C	86	21	0.8	1808	6	AX283629	AX283629 Sequence
C	87	21	0.8	2026	6	AR179702	AR179702 Sequence
C	88	21	0.8	3091	6	AX139124	AX139124 Sequence
C	89	21	0.8	4207	6	AX359930	AX359930 Sequence
C	90	21	0.8	4207	6	AX382144	AX382144 Sequence
C	91	21	0.8	4210	6	AX359931	AX359931 Sequence
C	92	21	0.8	4210	6	AX382145	AX382145 Sequence
C	93	21	0.8	6255	6	AX359939	AX359939 Sequence
C	94	21	0.8	6255	6	AX382153	AX382153 Sequence
C	95	21	0.8	9511	6	AX359960	AX359960 Sequence
C	96	21	0.8	9511	6	AX382174	AX382174 Sequence
C	97	21	0.8	37672	6	AX695374	AX695374 Sequence
C	98	21	0.8	137942	10	AL645745	AL645745 Mouse DNA
C	99	21	0.8	163245	2	AC013699	AC013699 Homo sapi
C	100	21	0.8	165940	2	AC109753	AC109753 Rattus no

ALIGNMENTS

Db	61	ACCCACGGGTCCGGAGTAGTGGTTGAGCTCGGCTGTTCTCCATTGTCAAGCAAGTCTATT	120
QY	121	CCAGATTGTTGAACCTTCTCTGGCCGCAATACAGAGAGAGACTTAAGCAAGCAAGG	180
Db	121	CCAGATTGTTGAACCTTCTCTGGCCGCAATACAGAGAGAGACTTAAGCAAGCAAGG	180
QY	181	GACCTACAGCCGCTGCGACATGGGCTGGTTAACTAGATTGTCGTCTTTCTGGGAGT	240
Db	181	GACCTACAGCCGCTGCGACATGGGCTGGTTAACTAGATTGTCGTCTTTCTGGGAGT	240
QY	241	ATTACTTAACAGCAAGCAAACTATCAGATGGAGAGACATGTGCCAAGGCTGAATT	300
Db	241	ATTACTTAACAGCAAGCAAACTATCAGATGGAGAGACATGTGCCAAGGCTGAATT	300
QY	301	ATCCACAAAGAAATGTTGSAATCCAAATGTATCATCTTCAATGGCTTGGGCCAAG	360
Db	301	ATCCACAAAGAAATGTTGSAATCCAAATGTATCATCTTCAATGGCTTGGGCCAAG	360
QY	361	CTCCAGTTATCATACCTTCTCTTTGGATGAGAAACGAGTAGGCTGTATGTTGAGCAAA	420
Db	361	CTCCAGTTATCATACCTTCTCTTTGGATGAGAAACGAGTAGGCTGTATGTTGAGCAAA	420
QY	421	GGATACACATTTTCATTTCGACCTGGGTAAATATCAAGGATTTTCAAAAGATTGTGGCC	480
Db	421	GGATACACATTTTCATTTCGACCTGGGTAAATATCAAGGATTTTCAAAAGATTGTGGCC	480
QY	481	AGTATCTTACACAGAGAGATGATGCAAGTGGCTGGAAAGACATCTCGAAAGATG	540
Db	481	AGTATCTTACACAGAGAGATGATGCAAGTGGCTGGAAAGACATCTCGAAAGATG	540
QY	541	TGCTTAATTTTCATCAAGGACTTAAGGCATATATCAGACTCATCTTGTACGCTGTGAAAC	600
Db	541	TGCTTAATTTTCATCAAGGACTTAAGGCATATATCAGACTCATCTTGTACGCTGTGAAAC	600

	RESULT 1			
AX207154	LOCUS	2709 bp	DNA	linear PAT 30-AUG-2001
DEFINITION	Sequence 1 from Patent WO0155455.			
ACCESSION	AX207154			
VERSION	AX207154.1	GI:15394943		
KEYWORDS	.			
SOURCE	Mus musculus (house mouse)			
ORGANISM	Mus musculus			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.			
TITLE	Jin,S.			
JOURNAL	Resistant sequences and uses thereof			
FEATURES	Patent: WO 0155455-A 1 02-VUQ-2001;			
	Millennium Pharmaceuticals, Inc. (US) ; Jin, Shengfang (US)			
source	Location/Qualifiers			
	1..2709			
	/organism="Mus musculus"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:10090"			
BASE COUNT	831 a 560 c 620 g 698 t			
ORIGIN				

Query Match	Similarity	Score	2709;	DB	6;	Length	2709;
Best Local	Similarity	100.0%;	Pred	2709;	0;	Mismatches	0;
Matches	2709;	Conservative	0;	Indels	0;	Gaps	0
QY	1	AATCTTTATTTTATTCGATGTTAAACAAGCTTACTATTCGATGCGACAGCTGAGGGGTGTG	60				
Db	1	AATCTTTATTTTATTCGATGTTAAACAAGCTTACTATTCGATGCGACAGCTGAGGGGTGTG	60				
QY	61	ACCCAGCGCTCGGGAGTAGTGGAGTGCCTCTCTCCCATTTGTCAGCGACTATTTT	120				

1021 AGGCGACGAGAGCTGTGTGATTAATGAGCAACATTCCTCAAGCTGTGATTTGGTC 1080

1081 AGTGGCAAGTCCAAATGGCATTTGACACTCATTTTGATGAACCTGGAGATGTATTCCTAT 1140

1081 AGTGCAGGTCCAAATGGCATTTGACACTCATTTTGATGAACCTGGAGATGTATTCCTAT 1140

1141 GAACCTTAAGATCCTTAATAAAATCCAGTTGTATATGAGAGTGTTCAGACTTCAGTAAAT 1200

PENSNKXHLQENKGRNRTHFERAPRSV"
BASE COUNT 786 a 518 c 576 g 650 t
ORIGIN
Query Match 92.6%; Score 2508; DB 6; Length 2530;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 193 CTGCACATGAGGCTGGTAACTAGATTGCTCTTTCTGGGAGTATTACTTACAGC 252
DB 9 CTGCACATGAGGCTGGTAACTAGATTGCTCTTTCTGGGAGTATTACTTACAGC 68
QY 253 AAGAGCAAACTATCAGAAATGGGAAGAACAAATGTGCCAAGGCTGAAATTAATCTCAAAAGA 312
DB 69 AAGAGCAAACTATCAGAAATGGGAAGAACAAATGTGCCAAGGCTGAAATTAATCTCAAAAGA 128
QY 313 AATGTGGAAATCCAAATGTGATCTTTCATATGCTTGGCCAAAGCTCCAGTTATCA 372
DB 129 AATGTGGAAATCCAAATGTGATCTTTCATATGCTTGGCCAAAGCTCCAGTTATCA 188
QY 373 TACCTTCTTTGGATGAGGAACGAGTAGGCTGTATGTTGGACAAAGATCAATAT 432
DB 189 TACCTTCTTTGGATGAGGAACGAGTAGGCTGTATGTTGGACAAAGATCAATAT 248
QY 433 TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGTGCCAGTATCTTAAC 492
DB 249 TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTGTGTGCCAGTATCTTAAC 308
QY 493 CAGAAGATGAAATGGAAGTGGCTGGAAAGACCTCGAAAGAAATGTCTAATTTCA 552
DB 309 CAGAAGATGAAATGGAAGTGGCTGGAAAGACCTCGAAAGAAATGTCTAATTTCA 368
QY 553 CAGGACTTAAAGCATATATATCAGACTCACTGTACGCTGTGAAACGGGGCTTTTGA 612
DB 369 CAGGACTTAAAGCATATATATCAGACTCACTGTACGCTGTGAAACGGGGCTTTTGA 428
QY 613 TCCAAATTTGCACTTGAATTAATGGAATGGAATCTCTGAGGACAAATATTTTAAAGCTGA 672
DB 429 TCCAAATTTGCACTTGAATTAATGGAATGGAATCTCTGAGGACAAATATTTTAAAGCTGA 488
QY 673 GAACTCAATTTTGAAGAGGGCGGTGGGAAGTCCATATGACCCATACCTGTGACAC 732
DB 489 GAACTCAATTTTGAAGAGGGCGGTGGGAAGTCCATATGACCCATACCTGTGACAC 548
QY 733 ATCCCTTTAAATAGATGAGAAATTAATCTGTGAACTGCAAGCTGATTTTATGGGGGAGA 792
DB 549 ATCCCTTTAAATAGATGAGAAATTAATCTGTGAACTGCAAGCTGATTTTATGGGGGAGA 608
QY 793 CTTTGCTATCTTCGAACTCTTGGGACCAACCAATTCAGACAGACAGCATGATTC 852
DB 609 CTTTGCTATCTTCGAACTCTTGGGACCAACCAATTCAGACAGACAGCATGATTC 668
QY 853 CAGGTGGCTCAATGATCCAAAGTTCATTTAGTCCCACTCATCTGAGAGTGAACAATCC 912
DB 669 CAGGTGGCTCAATGATCCAAAGTTCATTTAGTCCCACTCATCTGAGAGTGAACAATCC 728
QY 913 TGAAGATGACAAAGTATCTTTTCTTCGCTGAAATATGCAATATGATGAGAACACTCTG 972
DB 729 TGAAGATGACAAAGTATCTTTTCTTCGCTGAAATATGCAATATGATGAGAACACTCTG 788
QY 973 AAAAGTACTCAGCTAGATATGCTGATATGCAAGAAATGACTTTGAGGGGACAGAG 1032
DB 789 AAAAGTACTCAGCTAGATATGCTGATATGCAAGAAATGACTTTGAGGGGACAGAG 848
QY 1033 TCTGTGAATTAATGGAACAACATTCCTCAAGCTGCTGATTTGTCTAGTGCAGGTC 1092
DB 849 TCTGTGAATTAATGGAACAACATTCCTCAAGCTGCTGATTTGTCTAGTGCAGGTC 908
QY 1093 AAATGSCATTTGACACTCATTTTATGATGATGAGATGATTTCTAATGAACCTTTAAAGA 1152
DB 909 AAATGSCATTTGACACTCATTTTATGATGATGAGATGATTTCTAATGAACCTTTAAAGA 968
QY 1153 TCTTAAATAATCAAGTTTATATAGAGTGTTTACGACTTCCAGTAACTTTTCAAGGATTC 1212

DB 969 TCTTAAATAATCAAGTTTATATAGAGTGTTTACGACTTCCAGTAACTTTTCAAGGATTC 1028
QY 1213 AGCCGTGTATATGATATACATGATGATGAGAGGGGTCTTGGCCATATGGCCA 1272
DB 1029 AGCCGTGTATATGATATGATGATGAGAGGGGTCTTGGCCATATATGGCCA 1088
QY 1273 CAGGATGGAACCAACTATCAATGGGTCTTATCAAGGAAGATCCCTATCAACGGCC 1332
DB 1089 CAGGATGGAACCAACTATCAATGGGTCTTATCAAGGAAGATCCCTATCAACGGCC 1148
QY 1333 AGGAATCTGTCCAGCAAAACATTTGTGTGTTTACCTTACAAAGAACCTTCTGATGA 1392
DB 1149 AGGAATCTGTCCAGCAAAACATTTGTGTGTTTACCTTACAAAGAACCTTCTGATGA 1208
QY 1393 TGTATAACCTTTGGAAGAGATCCAGCAATGATCAATCCAGTTTCTATGAACAA 1452
DB 1209 TGTATAACCTTTGGAAGAGATCCAGCAATGATCAATCCAGTTTCTATGAACAA 1268
QY 1453 TCGCCCAATAGTATCAAAACGAGATGAATTAATTAACAAATTTGTCTAGAACG 1512
DB 1269 TCGCCCAATAGTATCAAAACGAGATGAATTAATTAACAAATTTGTCTAGAACG 1328
QY 1513 AGTGAATGCAAGATGGAACAGTATGATTTATCGGAACAGATGTGGACCGT 1572
DB 1339 AGTGAATGCAAGATGGAACAGTATGATTTATCGGAACAGATGTGGACCGT 1388
QY 1573 TCTTAAAGTATTAATTCCTAAGAGACTTGTATGATTTAAGAAAGGTTCTGCTGA 1632
DB 1389 TCTTAAAGTATTAATTCCTAAGAGACTTGTATGATTTAAGAAAGGTTCTGCTGA 1448
QY 1633 AGAAATGACAGTTTTCGGGAACCGACTGCTAATTTACAAATGAGCTTTCATAGCA 1692
DB 1449 AGAAATGACAGTTTTCGGGAACCGACTGCTAATTTACAAATGAGCTTTCATAGCA 1508
QY 1693 GCAAACTATATATTTGTTCAAGGCTGGGGTGGCCAGCTCCCTTTACACCGGTGGA 1752
DB 1509 GCAAACTATATATTTGTTCAAGGCTGGGGTGGCCAGCTCCCTTTACACCGGTGGA 1568
QY 1753 TATTTAAGGGAAGAGGTGCTGAGTGTGCTGCGCCGAGACCTTACTGCTGTGGGA 1812
DB 1569 TATTTAAGGGAAGAGGTGCTGAGTGTGCTGCGCCGAGACCTTACTGCTGTGGGA 1628
QY 1813 TGGTTCGATGTTCTGCTATTTTCCCATCTGCAAGAGAGCGACAAAGATAT 1872
DB 1629 TGGTTCGATGTTCTGCTATTTTCCCATCTGCAAGAGAGCGACAAAGATAT 1688
QY 1873 AAGAAATGAGACCCACTGACTCTACTGTTCAAGCTTAACCAATGATATACCATGGCCA 1932
DB 1689 AAGAAATGAGACCCACTGACTCTACTGTTCAAGCTTAACCAATGATATACCATGGCCA 1748
QY 1933 CAGCCCTGAAGAGAAATCATCTATGCTGTAGAGAAATGATGACATTTTGGAAATGAG 1992
DB 1749 CAGCCCTGAAGAGAAATCATCTATGCTGTAGAGAAATGATGACATTTTGGAAATGAG 1808
QY 1993 TCCGAATGCGAGAGAGCGGTGCTATTTGCAATTCAGAGGCGAAATGAAGCGAAA 2052
DB 1809 TCCGAATGCGAGAGAGCGGTGCTATTTGCAATTCAGAGGCGAAATGAAGCGAAA 1868
QY 2053 AGAAGATCAGAGTGAATGATCATATCATACAGACAAATCAAGGCCCTTTCGCTACGAG 2112
DB 1869 AGAAGATCAGAGTGAATGATCATATCATACAGACAAATCAAGGCCCTTTCGCTACGAG 1928
QY 2113 TCTACACAGAGAGATTCAGGCAATTTACCTCTGCAATCGGTGGAACATGGGTTATACA 2172
DB 1929 TCTACACAGAGAGATTCAGGCAATTTACCTCTGCAATCGGTGGAACATGGGTTATACA 1988
QY 2173 AACTCTTCTTAAGTAACTCTGAAATGATTAACACAGAGATTTTGAAGAACTTCTTCA 2232
DB 1989 AACTCTTCTTAAGTAACTCTGAAATGATTAACACAGAGATTTTGAAGAACTTCTTCA 2048
QY 2233 TAAAGATGATGAGATGGCTCTAAGACAAAGAAATGTCCATATGACATGACCTGAG 2292

	Best Local Similarity	100.0%	Pred. No. 0;	Matches 2508;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0
Qy	193	CTGCAGCATGGGCTGGTTAACTAGAAATGTCGTCTTTCTGGGAGTATTACTTACAGC	252					
Db	9	CTGCAGCATGGGCTGGTTAACTAGAAATGTCGTCTTTCTGGGAGTATTACTTACAGC	68					
Qy	253	AAGAGCAAACTATCAGAAATGGGAAAGAACATATGTCCAAAGGCTGAATTATCCTACAAAGA	312					
Db	69	AAGAGCAAACTATCAGAAATGGGAAAGAACATATGTCCAAAGGCTGAATTATCCTACAAAGA	128					
Qy	313	AATGTGGAATCCAAACATGTGATCATCTTTCAAATGGGCTGGCCAAACAGCTCAGTTATCA	372					
Db	129	AATGTGGAATCCAAACATGTGATCATCTTTCAAATGGGCTGGCCAAACAGCTCAGTTATCA	188					
Qy	373	TACCTTCCTTTGGATGAGAACGGAGTAGGCTGATGTTGGAGCAAGATCATATTT	432					
Db	189	TACCTTCCTTTGGATGAGAACGGAGTAGGCTGATGTTGGAGCAAGATCATATTT	248					
Qy	433	TTCAATTTGACCTGGTTAAATCAGAAATTTCAAAAAGATTGTGTGGCCAGTATCTTAAAC	492					
Db	249	TTCAATTTGACCTGGTTAAATCAGAAATTTCAAAAAGATTGTGTGGCCAGTATCTTAAAC	308					
Qy	493	CAGAAAGATGAATGCAAGTGGGCTGGAAAAAGATCCTGAAAGAAATGCTAAATTTAT	552					
Db	309	CAGAAAGATGAATGCAAGTGGGCTGGAAAAAGATCCTGAAAGAAATGCTAAATTTAT	368					
Qy	553	CAAGTACTTAAGGCATATATATCAGACTCATCTTGTAGGCTGTGGAAACGGGGGCTTTTCA	612					
Db	369	CAAGTACTTAAGGCATATATATCAGACTCATCTTGTAGGCTGTGGAAACGGGGGCTTTTCA	428					
Qy	613	TCCAAATTTGCACCTACATTGAAATTTGGACATCATCTGAGGACAAATATTTTAAAGCTGA	672					
Db	429	TCCAAATTTGCACCTACATTGAAATTTGGACATCATCTGAGGACAAATATTTTAAAGCTGA	488					
Qy	673	GAACTCAATTTTGA AAAACGGGCGGTGGGAAAGATGCAATGACCTTAAGCTGTGACAGC	732					
Db	489	GAACTCAATTTTGA AAAACGGGCGGTGGGAAAGATGCAATGACCTTAAGCTGTGACAGC	548					
Qy	733	ATCCCTTTTAATAGATGAGAAATTAATATCTCGAAATGCAAGCTGATTTTATGGGGCGAGA	792					
Db	549	ATCCCTTTTAATAGATGAGAAATTAATATCTCGAAATGCAAGCTGATTTTATGGGGCGAGA	608					
Qy	793	CTTTGCTATCTTCCGAACCTTTGGGCAACCAACCAATCAGGACAGATGATTC	852					
Db	609	CTTTGCTATCTTCCGAACCTTTGGGCAACCAACCAATCAGGACAGATGATTC	668					
Qy	853	CAGTGGCTCAATGATCCAAAGTTCAATTAATGTCCTCATCTTCAGAGATGCAATCC	912					
Db	669	CAGTGGCTCAATGATCCAAAGTTCAATTAATGTCCTCATCTTCAGAGATGCAATCC	728					
Qy	913	TGAAGATGACAAAGATTAATCTTTCTTCGCGGAAATATGCAATATGAGAGAACTCTGG	972					
Db	729	TGAAGATGACAAAGATTAATCTTTCTTCGCGGAAATATGCAATATGAGAGAACTCTGG	788					
Qy	973	AAAAGCTACTCAGCTGAAATAGGTGCAGATATGCAAGATGACTTTGGAGGGCACAGAG	1032					
Db	789	AAAAGCTACTCAGCTGAAATAGGTGCAGATATGCAAGATGACTTTGGAGGGCACAGAG	848					
Qy	1033	TCTGTGTAATAATATGACAACTTCTCCAAAGCTGTCTGATTTGGTCTCAGTGCAGGTCC	1092					
Db	849	TCTGTGTAATAATATGACAACTTCTCCAAAGCTGTCTGATTTGGTCTCAGTGCAGGTCC	908					
Qy	1093	AAATGGCATGGACACTCAITTTGATGAAATGCAAGATGATTTCCATAATGAACCTTAAAGA	1152					
Db	909	AAATGGCATGGACACTCAITTTGATGAAATGCAAGATGATTTCCATAATGAACCTTAAAGA	968					
Qy	1153	TCCTAAATAATCCAGTTGTATATGGAAGTGTTCAGACTTCAGATTAACATTTTCAAGGATC	1212					
Db	969	TCCTAAATAATCCAGTTGTATATGGAAGTGTTCAGACTTCAGATTAACATTTTCAAGGATC	1028					
Qy	1213	AGCCGTGTATATATATGCAATGATGATATGGAAGGATGTTCTTGGTCCATATGCCCC	1272					

GDGSKTKEMSNMTSPSÖKYWRIDEMQLINHPNLTMTDEFCEQWKRDRKÖRRÖRPHGT
PENSNMKWHLOENKGRNRRTHEFERARPSV
BASE COUNT 786 a 518 c 576 g 650 t
ORIGIN

Query Match 92.6%; Score 2508; DB 9; Length 2530;
Beet Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 193 CTGCAGCATGGGCTGTTAACTAGAGATGTCTGTCTTTCTGGGGAGTATTACTTAACG 252
DB 9 CTGCAGCATGGGCTGTTAACTAGAGATGTCTGTCTTTCTGGGGAGTATTACTTAACG 68
OY 253 AAGAGCAAACTATCAGAAATGGGAGAACATGTGCCAAGGCTGAATTTCTCTCAAAAGA 312
DB 69 AAGAGCAAACTATCAGAAATGGGAGAACATGTGCCAAGGCTGAATTTCTCTCAAAAGA 128
OY 313 AATGTGGAATTCACATATGTATCACTTTCAATGGCTGGCCAAACGCTCAAGTTATCA 372
DB 129 AATGTGGAATTCACATATGTATCACTTTCAATGGCTGGCCAAACGCTCAAGTTATCA 188
OY 373 TACCTTCCCTTTGGATGAGAAACGAGATAGGCTGTATGTGGAGCAAGATCAATAT 432
DB 189 TACCTTCCCTTTGGATGAGAAACGAGATAGGCTGTATGTGGAGCAAGATCAATAT 248
OY 433 TTCAATTCGACCTGTGTTAATATCAAGATTTTCAAAAGATGTGTGCCAGTATCTTAAC 492
DB 249 TTCAATTCGACCTGTGTTAATATCAAGATTTTCAAAAGATGTGTGCCAGTATCTTAAC 308
OY 493 CAGAAAGATGATGATGCAAGTGGGCTGGAAGAACATCTCTGAAGATGTCTAATTTCA 552
DB 309 CAGAAAGATGATGATGCAAGTGGGCTGGAAGAACATCTCTGAAGATGTCTAATTTCA 368
OY 553 CAGGATCTTAAGGATATATACAGATCACTTGTAGCCCTGTGGAAGGGGGCTTTCA 612
DB 369 CAGGATCTTAAGGATATATACAGATCACTTGTAGCCCTGTGGAAGGGGGCTTTCA 428
OY 613 TCCAAATTTGCACTATGAAATTTGACATCATCTGAGGACATATTTTAAAGCTGGA 672
DB 429 TCCAAATTTGCACTATGAAATTTGACATCATCTGAGGACATATTTTAAAGCTGGA 488
OY 673 GAATCACAATTTGAAAAAGCGCGTGGGAGAGATCATATGACCTTAAGCTGACAGC 732
DB 489 GAATCACAATTTGAAAAAGCGCGTGGGAGAGATCATATGACCTTAAGCTGACAGC 548
OY 733 ATCCCTTTAATAGATGAGAAATTAATACTGGAAGTGCAGGTGATTTTAAAGGGCGGA 792
DB 549 ATCCCTTTAATAGATGAGAAATTAATACTGGAAGTGCAGGTGATTTTAAAGGGCGGA 608
OY 793 CTTTGTCTATCTTCCGAATCTTTGGGACCAACCCATCAAGACAGAGCATGATTC 852
DB 609 CTTTGTCTATCTTCCGAATCTTTGGGACCAACCCATCAAGACAGAGCATGATTC 668
OY 853 CAGGTGCTCAATGATCAAAAGTTCAATAGTCCCACTCATCTCAGAGATGACAAATCC 912
DB 669 CAGGTGCTCAATGATCAAAAGTTCAATAGTCCCACTCATCTCAGAGATGACAAATCC 728
OY 913 TGAAGATGACAAAGATATCTTTTCTCCGTGGAATTCATATGATGGAACACTCTGG 972
DB 729 TGAAGATGACAAAGATATCTTTTCTCCGTGGAATTCATATGATGGAACACTCTGG 788
OY 973 AAAAGCTACGCTAGATAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1032
DB 789 AAAAGCTACGCTAGATAGATGATGATGATGATGATGATGATGATGATGATGATGATG 848
OY 1033 TCTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1092
DB 849 TCTGTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 908
OY 1093 AAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1152
DB 909 AAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 968

OY 1153 TCTTAAATATCAGATGATATGAGATGTTTACGATCTTCCAGTAACTTTTCAAGGATC 1212
DB 969 TCTTAAATATCAGATGATATGAGATGTTTACGATCTTCCAGTAACTTTTCAAGGATC 1028
OY 1213 AGCCGTGTATGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1272
DB 1029 AGCCGTGTATGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1088
OY 1273 CAGGATGAGACCAACTATCAATGGGTGCTTATCAAGGAAGATCCCTTATCCAGGCTC 1332
DB 1089 CAGGATGAGACCAACTATCAATGGGTGCTTATCAAGGAAGATCCCTTATCCAGGCTC 1148
OY 1333 AGGAATCTGTCCAGCAAAACATTTGATGTTTGAATCTTCAAAAGACCTTCTGATGA 1392
DB 1149 AGGAATCTGTCCAGCAAAACATTTGATGTTTGAATCTTCAAAAGACCTTCTGATGA 1208
OY 1393 TGTATTAACCTTTTCAAGAGATGATCCAGCATGATGATGATGATGATGATGATGATG 1452
DB 1209 TGTATTAACCTTTTCAAGAGATGATCCAGCATGATGATGATGATGATGATGATGATG 1268
OY 1453 TCGCCCAATAGTATCAAAACGAGATGAATTAATTAATTAATTAATTAATTAATTAAT 1512
DB 1269 TCGCCCAATAGTATCAAAACGAGATGAATTAATTAATTAATTAATTAATTAATTAAT 1328
OY 1513 AGTGAATGACAGAGATGACAGATGATGATGATGATGATGATGATGATGATGATGATG 1572
DB 1329 AGTGAATGACAGAGATGACAGATGATGATGATGATGATGATGATGATGATGATGATG 1388
OY 1573 TCTTAAAGTATGTTCAATTTCTTAAGAGACTTGTATGATTTAGAAAGTTCTGTGGA 1632
DB 1389 TCTTAAAGTATGTTCAATTTCTTAAGAGACTTGTATGATTTAGAAAGTTCTGTGGA 1448
OY 1633 AGAATGACAGTTTGGGGACCGATGATTTTCAAGATGAGATGATGATGATGATGATG 1692
DB 1449 AGAATGACAGTTTGGGGACCGATGATTTTCAAGATGAGATGATGATGATGATGATG 1508
OY 1693 GCAACAACTATATATTTGTTCAACGCTGGGTTGCCAGCTCCCTTACACCGGTGGA 1752
DB 1509 GCAACAACTATATATTTGTTCAACGCTGGGTTGCCAGCTCCCTTACACCGGTGGA 1568
OY 1753 TATTTACGGGAAGCGTGTGCTGATGTTGCTGCGCCGAGACCTTACTGTCTGGGA 1812
DB 1569 TATTTACGGGAAGCGTGTGCTGATGTTGCTGCGCCGAGACCTTACTGTCTGGGA 1628
OY 1813 TGTGTTGCAATGTTCTGCTATTTTCCACCTGCAAAAGACGCAAGAGCAAGATAT 1872
DB 1629 TGTGTTGCAATGTTCTGCTATTTTCCACCTGCAAAAGACGCAAGAGCAAGATAT 1688
OY 1873 AAGAAATGAGACCCACTGACTCACTGTTCACTTACACCATGATATGACATGAGCCA 1932
DB 1689 AAGAAATGAGACCCACTGACTCACTGTTCACTTACACCATGATATGACATGAGCCA 1748
OY 1933 CAGCCTTGAAGAGAAATCATATGATGATGATGATGATGATGATGATGATGATGATG 1992
DB 1749 CAGCCTTGAAGAGAAATCATATGATGATGATGATGATGATGATGATGATGATGATG 1808
OY 1993 TCCGAATGCGAGAGAGCGCTGTCTATTTGGCAATTCACAGGCGGAATGAAAGCGAAA 2052
DB 1809 TCCGAATGCGAGAGAGCGCTGTCTATTTGGCAATTTCCAGAGGGAATGAAAGCGAAA 1868
OY 2053 AGAAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2112
DB 1869 AGAAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1928
OY 2113 TCTTCAACAGAGATTCAGGCAATTAATCTCTGATGATGATGATGATGATGATGATGATG 2172
DB 1929 TCTTCAACAGAGATTCAGGCAATTAATCTCTGATGATGATGATGATGATGATGATGATG 1988
OY 2173 AACTCTTAAAGTAAACCTGGAAGATGATGATGATGATGATGATGATGATGATGATG 2232
DB 1989 AACTCTTAAAGTAAACCTGGAAGATGATGATGATGATGATGATGATGATGATGATG 2048
OY 2233 TAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2292

Db	2049	TAAATATGATGATGGAGATGGCTCTTAAGCCAAAGAAATGTCCATTAGCATGACCTTAG	2108
QY	2293	CCAGAAAGGTCGTGTACAGAGACTTCATGACGCTCATCAACCCCAATCTTCAACAGAT	2353
Db	2109	CCAGAAAGGCTGTGTACAGAGACTTCATGACGCTCATCAACCCCAATCTTCAACAGAT	2168
QY	2353	GGATGAGTTCGTGTGAACAAGTTTGGAAAAAGGACCGAAAAACAAGTGGCAAAAGCCACGG	2412
Db	2169	GGATGAGTTCGTGTGAACAAGTTTGGAAAAAGGACCGAAAAACAAGTGGCAAAAGCCACGG	2228
QY	2413	ACATACCCACAGGGAACAGTAACAATTGGAGAAGCACTTACAAGAAAAATTAAGAAAGGTAGAA	2472
Db	2229	ACATACCCACAGGGAACAGTAACAATTGGAGAAGCACTTACAAGAAAAATTAAGAAAGGTAGAA	2288
QY	2473	CAGAGAGACCCACGAATTTGAGAGGGACCCAGAGAGTCTGAGCTGCAATTACCTTAGA	2532
Db	2289	CAGAGAGACCCACGAATTTGAGAGGGACCCAGAGAGTGTCTGAGCTGCAATTACCTTAGA	2348
QY	2533	AACCTCAAAACAAGTAGAAACCTTGCTTGAACAATAACTGAGAAAAACAATGCAATTACAT	2592
Db	2349	AACCTCAAAACAAGTAGAAACCTTGCTTGAACAATAACTGAGAAAAACAATGCAATTACAT	2408
QY	2593	GAACTTTTTCATGGCATTAATGATGGATGTTTACAATGTTGGGAAATTCAGCTGAGTCCA	2653
Db	2409	GAACTTTTTCATGGCATTAATGATGTTTACAATGTTGGGAAATTCAGCTGAGTCCA	2468
QY	2653	CCAATTATAAATTAATCCATGAGTAACTTTCTTAATGAGGCTTTT	2700
Db	2469	CCAATTATAAATTAATCCATGAGTAACTTTCTTAATGAGGCTTTT	2516

RESULT 5				
LOCUS	G31703	2530 bp	DNA	linear
DEFINITION	GWSMS1973 Eric D. Green Homo sapiens STS genomic, sequence tagged site.			
ACCESSION	G31703			
VERSION	G31703.1	GI:1916428		
KEYWORDS	STS.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 2530)			
AUTHORS	Bouffard,G.G., Iyer,L.M., Idol,J.R., Braden,V.V., Cunningham,A.F., Weintraub,L.A., Mohr-Tidwell,R.M., Peluso,D.C., Fulton,R.S., Leckie,M.P. and Green,E.D.			
	A collection of 1814 human chromosome 7-specific STS			
TITLE	Genome Res. 7 (1), 59-64 (1997)			
JOURNAL	97189344			
MEDLINE	9037602			
PubMed	2 (bases 1 to 2530)			
REFERENCE	Green,E.D.			
AUTHORS	Human chromosome 7 STSs (1997)			
TITLE	Unpublished (1997)			
JOURNAL	Synonyms: SEMITI			
COMMENT				

Contact: Eric D. Green
 Genome Technology Branch
 National Human Genome Research Institute/NIH
 49 Convent Dr., MSC4431, Bldg. 49, Rm. 2A08, Bethesda, MD 20892
 Tel: 3014020201
 Fax: 3014024735
 Email: egreen@hgrl.nih.gov
 Primer A: CCTGTGAGTCATTGACACAG
 Primer B: GGATGTCTGTGGCTTTGCCG
 STS size: 230
 PCR Profile:
 Presoak: 0 degrees C for 0.00 minute(s)
 Denaturation: 92 degrees C for 0.17 minute(s)
 Annealing: 55 degrees C for 1.00 minute(s)
 Polymerization: 72 degrees C for 1.00 minute(s)

PCR Cycles:	35
Thermal Cycler:	PerkinElmer 9600
Protocol:	
Template:	30-100 ng
Primer:	each 1 μ M
dNTPs:	each 200 μ M
Taq Polymerase:	0.05 units/ μ l
Total Vol:	10 μ l
Buffer:	
MgCl ₂ :	1.5 mM
KCl:	100 mM
Tris-HCl:	10 mM
NH ₄ Cl:	5 mM
pH:	8.6

This STS has been incorporated into the NHGRI chromosome 7 physical map, but was developed by another investigator. See GenBank record: U26081. For additional information about the NHGRI chromosome 7 mapping project, see <http://www.nhgri.nih.gov/DIR/GTE/CHR7>. Also see Genomics 11:548-64 (1991) [MUID=92128937].

FEATURES	
source	location/Qualifiers
	1..2530
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/map="7"
	/clone_lib="Eric D. Green"
gene	1..2530
	/gene="SEMI11"
STS	2006..2235
	/gene="SEMI11"
primer_bind	2006..2026
	/gene="SEMI11"
primer_bind	complement(2215..2235)
BASE COUNT	786 a 518 c 576 g 650 t
ORIGIN	

Query Match	Best Local Similarity	92.6%; Score 2508; DB 11; Length 2530;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
Qy	193	CTGCAGCATGGGCTGGTTAACTAGATTGTCGTCTTTTCTGGGAGTATTACTTACAGC 252
Db	9	CTGCAGCATGGGCTGGTTAACTAGATTGTCGTCTTTTCTGGGAGTATTACTTACAGC 68
Qy	253	AAGGCAAACTATATAGAAATGGGAAGAACATGTGCCAAGCTGAATTTATCTTACAAGA 312
Db	69	AAGGCAAACTATATAGAAATGGGAAGAACATGTGCCAAGCTGAATTTATCTTACAAGA 128
Qy	313	AATGTGGAAATCCAAACATGTGATCACTTTCAATGGCTTGCGCCAAAGCTCCAGTTATCA 372
Db	129	AATGTGGAAATCCAAACATGTGATCACTTTCAATGGCTTGCGCCAAAGCTCCAGTTATCA 188
Qy	373	TACCTTCCTTTTGGATGAGGAACGAGTAGCGTGTATGTTGGAGCAAAAGATCACATATT 432
Db	189	TACCTTCCTTTTGGATGAGGAACGAGTAGCGTGTATGTTGGAGCAAAAGATCACATATT 248
Qy	433	TTCAATTGACCTGGTTAAATATCAAGATTTTTCAAAATAATTGTGTGGCCAGTATCTTAC 492
Db	249	TTCAATTGACCTGGTTAAATATCAAGATTTTTCAAAATAATTGTGTGGCCAGTATCTTAC 308
Qy	493	CAGAAAGATGAATGCAAGTGGCTGGAAAAAGACATCTGAAAGAAATGTCTAATTTCAT 552
Db	309	CAGAAAGATGAATGCAAGTGGCTGGAAAAAGACATCTGAAAGAAATGTCTAATTTCAT 368
Qy	553	CAAGTACTTAAGGCAATATAATACACTCATTTGTACGCTGTGGAAACGGGGGCTTTTCA 612
Db	369	CAAGTACTTAAGGCAATATAATACACTCATTTGTACGCTGTGGAAACGGGGGCTTTTCA 428
Qy	613	TCCAATTTGCACCTACATTGAAATTGACATCATCTGGAAGACAAATTTTTTAAAGCTGGA 672

SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2601)
AUTHORS Goodman, C.S., Kolodkin, A.L., Matthes, D., Bentley, D.R. and O'Connor, T.
TITLE Semaphorin gene family
JOURNAL Patent: US 5807826-A 53 15-SEP-1998;
FEATURES Location/Qualifiers
source 1..2601
BASE COUNT 809 a 533 c 593 g 666 t
ORIGIN
Query Match 92.6%; Score 2508; DB 6; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 193 CTGACGATGGGCTGGTAACTAGATGTTGTCCTTTCTGGGAGATTAATTACAGC 252
DB 9 CTGACGATGGGCTGGTAACTAGATGTTGTCCTTTCTGGGAGATTAATTACAGC 68
QY 253 AAGAGAACTATCAGAAATGGGAGAAACAATGTCAGGCTGAATTTCTTACAAAGA 312
DB 69 AAGAGAACTATCAGAAATGGGAGAAACAATGTCAGGCTGAATTTCTTACAAAGA 128
QY 313 AATGTTGAATCCAAATGTCATCTTTCATGAGCTTGGCCAAAGCTCCAGTTATCA 372
DB 129 AATGTTGAATCCAAATGTCATCTTTCATGAGCTTGGCCAAAGCTCCAGTTATCA 188
QY 373 TACCTTCTTTGGATGAGGAAACGAGTAGCTGTATGTTGGAGCAAGATCAATATT 432
DB 189 TACCTTCTTTGGATGAGGAAACGAGTAGCTGTATGTTGGAGCAAGATCAATATT 248
QY 433 TTCATTGCACTGGTAAATATCAAGATTTTCAAAGATTTGTCGCCAGTATCTTACAC 492
DB 249 TTCATTGCACTGGTAAATATCAAGATTTTCAAAGATTTGTCGCCAGTATCTTACAC 308
QY 493 CAGAGAGATGATGCAATGGGCTGGAAAGACATCTGAAAGAAATGTCTAATTTCAAT 552
DB 309 CAGAGAGATGATGCAATGGGCTGGAAAGACATCTGAAAGAAATGTCTAATTTCAAT 368
QY 553 CAGGTAATTAAGCATATATCACTCACTTTGAGCCCTGTGGAACGGGGCTTTTCA 612
DB 369 CAGGTAATTAAGCATATATCACTCACTTTGAGCCCTGTGGAACGGGGCTTTTCA 428
QY 613 TCCAAATTTGCATATGAAATTGGAATCATCTGAGACAAATTTTAAAGTGA 672
DB 429 TCCAAATTTGCATATGAAATTGGAATCATCTGAGACAAATTTTAAAGTGA 488
QY 673 GAATCACAATTTGAAAGCGCGTGGAGAGTCCATATGACCTTAAGCTGTCAGAC 732
DB 489 GAATCACAATTTGAAAGCGCGTGGAGAGTCCATATGACCTTAAGCTGTCAGAC 548
QY 733 ATCCCTTTAATAGATGAGAAATTATATCTGTGAAATGTCAGCTGATTTATGGGGGAGA 792
DB 549 ATCCCTTTAATAGATGAGAAATTATATCTGTGAAATGTCAGCTGATTTATGGGGGAGA 608
QY 793 CTTTGTATCTTCCGAACCTTGGGACACACCCCAATCAGAGACAGACAGATGATTC 852
DB 609 CTTTGTATCTTCCGAACCTTGGGACACACCCCAATCAGAGACAGACAGATGATTC 668
QY 853 CAGGTGCTCATGATCCAAAGTCTTATGAGCCCACTCATCTCAGAGATGACAAATCC 912
DB 669 CAGGTGCTCATGATCCAAAGTCTTATGAGCCCACTCATCTCAGAGATGACAAATCC 728
QY 913 TGAAGATGCAAAAGTATCTTTTCTTCCGTGAAAATGCAATGATGAGAACACTCTGG 972
DB 729 TGAAGATGCAAAAGTATCTTTTCTTCCGTGAAAATGCAATGATGAGAACACTCTGG 788
QY 973 AAAACTCTACGCTAGAAATGTCATGATGCAAGAAATGATCTTTGGAGGCAAGAAG 1032
DB 789 AAAACTCTACGCTAGAAATGTCATGATGCAAGAAATGATCTTTGGAGGCAAGAAG 848

QY 1033 TCTGTGATATAATGAGACAACTTCTCAAAAGCTGCTGATTTGCTCAGTCCAGGTCC 1092
DB 849 TCTGTGATATAATGAGACAACTTCTCAAAAGCTGCTGATTTGCTCAGTCCAGGTCC 908
QY 1093 AATGCAATGACACTCATTTTGTGAACCTGAGAGATGATTTCTTAATGAATTTAAGA 1152
DB 909 AATGCAATGACACTCATTTTGTGAACCTGAGAGATGATTTCTTAATGAATTTAAGA 968
QY 1153 TCCTAATAATCCAGTGTATATGAGGTGTTAAGCACTTCCAGTACATTTTCAAGGATC 1212
DB 969 TCCTAATAATCCAGTGTATATGAGGTGTTAAGCACTTCCAGTACATTTTCAAGGATC 1028
QY 1213 ACCCGTGTATGTATAGCATGAGATGATGAGAGGGGTGTTCTTGTGCTCATATGCCA 1272
DB 1029 ACCCGTGTATGTATAGCATGAGATGATGAGAGGGGTGTTCTTGTGCTCATATGCCA 1088
QY 1273 CAGGATGAGCCCAATCTATCAATGGGTGCTTATTAAGAAAGATCCCTATCCAGGCC 1332
DB 1089 CAGGATGAGCCCAATCTATCAATGGGTGCTTATTAAGAAAGATCCCTATCCAGGCC 1148
QY 1333 AGGAATGTCAGCAAAAATTTGGTGTGTTGACTCTAACAAGACCTTCCATGATGA 1392
DB 1149 AGGAATGTCAGCAAAAATTTGGTGTGTTGACTCTAACAAGACCTTCCATGATGA 1208
QY 1393 TGTATTAACCTTTGCAAGAGTCAATCCAGCATGTACATCCAGTGTTCATGAACA 1452
DB 1209 TGTATTAACCTTTGCAAGAGTCAATCCAGCATGTACATCCAGTGTTCATGAACA 1268
QY 1453 TCGCCCAATAGTATCAAAAAGGATGTAAATTAATTAATTTACAAATTTGCTGACCG 1512
DB 1269 TCGCCCAATAGTATCAAAAAGGATGTAAATTAATTAATTTACAAATTTGCTGACCG 1328
QY 1513 AGTGATGAGAAATGAGACAGTATGATGTTATGTTATCGGAACAGATGTTGGAGCGT 1572
DB 1329 AGTGATGAGAAATGAGACAGTATGATGTTATGTTATCGGAACAGATGTTGGAGCGT 1388
QY 1573 TCTTAAAGTATGTTCAATTCCTAAGAGACTTGTATGATTTAAGAGGTTCTGCTGA 1632
DB 1389 TCTTAAAGTATGTTCAATTCCTAAGAGACTTGTATGATTTAAGAGGTTCTGCTGA 1448
QY 1633 AGAAATGACAGTTTTCGGGAAACGACCTGCTATTTACCAATGAGCTTCCACTAACA 1692
DB 1449 AGAAATGACAGTTTTCGGGAAACGACCTGCTATTTACCAATGAGCTTCCACTAACA 1508
QY 1693 GCAACAATATATATGTTGTTCAACGGCTGGGTTGCCAGCTCCCTTACACGGGTGA 1752
DB 1509 GCAACAATATATATGTTGTTCAACGGCTGGGTTGCCAGCTCCCTTACACGGGTGA 1568
QY 1753 TATTTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGACCTTACTGTGGGA 1812
DB 1569 TATTTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGACCTTACTGTGGGA 1628
QY 1813 TGTGTTGCAATGTTTGGCTATTTTCCACTGCAAAAGAGACGACAAAGCAAGATAT 1872
DB 1629 TGTGTTGCAATGTTTGGCTATTTTCCACTGCAAAAGAGACGACAAAGCAAGATAT 1688
QY 1873 AAGAAATGAGACCCACTGACCTACTGTTCACTTAACCATGATATACATGAGCCA 1932
DB 1689 AAGAAATGAGACCCACTGACCTACTGTTCACTTAACCATGATATACATGAGCCA 1748
QY 1933 CAGCCCTGAAGAGAAATCTATGTTGTAAGAAATGATGACATTTTGGAAATGAG 1992
DB 1749 CAGCCCTGAAGAGAAATCTATGTTGTAAGAAATGATGACATTTTGGAAATGAG 1808
QY 1993 TCCGAAGTCCAGAGAGCGCTGTATTTGGCAATTTCAAGAGCGAAATGAAAGCGAAA 2052
DB 1809 TCCGAAGTCCAGAGAGCGCTGTATTTGGCAATTTCAAGAGCGAAATGAAAGCGAAA 1868
QY 2053 AGAAGAGATCAGAGTATGATCATATCATGAGGACAGATCAAGGCTTCTGCTACGTAG 2112
DB 1869 AGAAGAGATCAGAGTATGATCATATCATGAGGACAGATCAAGGCTTCTGCTACGTAG 1928

OY	2113	TCACAA CAGAGAGATT CAGCAATTA CTCTGCATCGGATGGAAATGGGTTCA TACA	2172
Db	1929	TCACAA CAGAGAGATT CAGCAATTA CTCTGCATCGGATGGAAATGGGTTCA TACA	1988
OY	2173	AAC TCTTCTTAAGGTAA CCGTGAAGTCATTTGACA CAGAGCATTTGGAAAGAACTCTTCA	2232
Db	1989	AAC TCTTCTTAAGGTAA CCGTGAAGTCATTTGACA CAGAGCATTTGGAAAGAACTCTTCA	2048
OY	2233	TAAAGATGATGATGGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACTAG	2292
Db	2049	TAAAGATGATGATGGAGATGGCTCTAAGACCAAGAAATGTCCAATAGCATGACACTAG	2108
OY	2293	CCAGAGGTC TGGTACAGAGA CTTCATGACACTATCAACCA CCGCAATCTCAACAGAT	2352
Db	2109	CCAGAGGTC TGGTACAGAGA CTTCATGACACTATCAACCA CCGCAATCTCAACAGAT	2168
OY	2353	GGATGAGTTC TGTGACAAAGTTTGGAAAGGGA CCGAAACMA CGTGGCCAAAGCCAGG	2412
Db	2169	GGATGAGTTC TGTGACAAAGTTTGGAAAGGGA CCGAAACMA CGTGGCCAAAGCCAGG	2228
OY	2413	ACATPACCC CAGGGAACGCTAACAAATGGAGCACTTACAGAAAAATAGAAAGGTAGAA	2472
Db	2229	ACATPACCC CAGGGAACGCTAACAAATGGAGCACTTACAGAAAAATAGAAAGGTAGAA	2288
OY	2473	CAGAGGACCA CCAAGATTGAGAGGGCA CCGAGAGATGCTGAGCTGATTA CCTCTAGA	2532
Db	2289	CAGAGGACCA CCAAGATTGAGAGGGCA CCGAGAGATGCTGAGCTGATTA CCTCTAGA	2348
OY	2533	AACCTCAAA CAAAGTAGAACTGCTGACATACATTAACGTGAAAAACAATGCATATATCAT	2592
Db	2349	AACCTCAAA CAAAGTAGAACTGCTGACATACATTAACGTGAAAAACAATGCATATATCAT	2408
OY	2593	GAACTTTTTCATGCGCATTAATGAGATGTTTACAATGATGGAAATTCA GCTGAGTTCCA	2652
Db	2409	GAACTTTTTCATGCGCATTAATGAGATGTTTACAATGATGGAAATTCA GCTGAGTTCCA	2468
OY	2653	CCAATTATTAATTAAATCCATGAATTA CTTCCATATAGGCTTTTTTTT	2700
Db	2469	CCAATTATTAATTAAATCCATGAATTA CTTCCATATAGGCTTTTTTTT	2516

QY	313	TATGTGGAAATCCAA	CATGTGATCACTTTGAATGGCTTGGCCAA	CAGCTCAGTATCA	372
Db	129	AATGTGAAATCCAA	CAATGTGATCACTTTGAATGGCTTGGCCAA	CAGCTCAGTATCA	188
QY	373	TACCTTCCTTTGGAT	TGAGAAACGAGTAGAGCTGATTTGGAGCAAAAGATTCATATTT	432	
Db	169	TACCTTCCTTTGGAT	TGAGAAACGAGTAGAGCTGATTTGGAGCAAAAGATTCATATTT	248	
QY	433	TTCAATTCGACCTG	TTAATATCAGAGATTTTCAAAAGATTGTGTGGCCAGTATCTTACAC	492	
Db	249	TTCAATTCGACCTG	TTAATATCAGAGATTTTCAAAAGATTGTGTGGCCAGTATCTTACAC	308	
QY	493	CAGAAAGATGAATCA	TGAGTGGGCTGGAAAAAGACATCCTGAAAGAAATGGCTAATTTTCAT	552	
Db	309	CAGAAAGATGAATCA	TGAGTGGGCTGGAAAAAGACATCCTGAAAGAAATGGCTAATTTTCAT	368	
QY	553	CAAGGTACTTAAGG	CAATATATCAGACTCACCTTGTACGCTGTGGAAACGGGGCTTTTCA	612	
Db	369	CAAGGTACTTAAGG	CAATATATCAGACTCACCTTGTACGCTGTGGAAACGGGGCTTTTCA	428	
QY	613	TTCAATTTGCACTTA	CTTAATTTGGAATTTGGAATCAATCCTGAGAGCAATATTTTAAAGCTGGA	672	
Db	429	TTCAATTTGCACTTA	CTTAATTTGGAATTTGGAATCAATCCTGAGAGCAATATTTTAAAGCTGGA	488	
QY	673	GAACATCAATTTTGA	AAAAACGGCCGTGGAGAAAGTCCATATGACCTTAAGCTGTGACAGC	732	
Db	489	GAACATCAATTTTGA	AAAAACGGCCGTGGAGAAAGTCCATATGACCTTAAGCTGTGACAGC	548	
QY	733	ATCCCTTTTAAATAG	TGAGAAATTAATCTGTGAACTGACAGCTGATTTTATGGGGCGAGA	792	
Db	549	ATCCCTTTTAAATAG	TGAGAAATTAATCTGTGAACTGACAGCTGATTTTATGGGGCGAGA	608	
QY	793	CTTTGCTATCTTCCG	AATCTTTGGGCAACACACCCCAATCAGGAACAGAGCAAGCATGATTC	852	
Db	609	CTTTGCTATCTTCCG	AATCTTTGGGCAACACACCCCAATCAGGAACAGAGCAAGCATGATTC	668	
QY	853	CAGGTGGCTCATGAT	CCAAAGTTCAATTAAGTCCCACTCATCTCAGAGATGACAAATCC	912	
Db	669	CAGGTGGCTCATGAT	CCAAAGTTCAATTAAGTCCCACTCATCTCAGAGATGACAAATCC	728	
QY	913	TGAAATGACAAAGTA	TATCTTTTCTTCCTGGAAAAATGCAATAGATGAGAAACATCTGTGG	972	
Db	729	TGAAATGACAAAGTA	TATCTTTTCTTCCTGGAAAAATGCAATAGATGAGAAACATCTGTGG	788	
QY	973	AAAAAGCTACTACG	CTAGATAGGTCAGATATGCAAGATATGACTTTGGAGGGCACAGAAAG	1032	
Db	789	AAAAAGCTACTACG	CTAGATAGGTCAGATATGCAAGATATGACTTTGGAGGGCACAGAAAG	848	
QY	1033	TCTGTGTAAATTA	TGACACATTTCTCAAAAGCTGTGTGATTTTGTCTCAGTCCAGGTCC	1092	
Db	849	TCTGTGTAAATTA	TGACACATTTCTCAAAAGCTGTGTGATTTTGTCTCAGTCCAGGTCC	908	
QY	1093	AAATGGCATTTGACA	CTCATTTTGTGATGAACTGACAGAGATATTTCTTAATGAACCTTTAAAGA	1152	
Db	909	AAATGGCATTTGACA	CTCATTTTGTGATGAACTGACAGAGATATTTCTTAATGAACCTTTAAAGA	968	
QY	1153	TCTTAAAAAATCAG	TTGTATATGAGAGTGTTTAGACATTCACAGTAACATTTTCAAGGATC	1212	
Db	969	TCTTAAAAAATCAG	TTGTATATGAGAGTGTTTAGACATTCACAGTAACATTTTCAAGGATC	1028	
QY	1213	AGCGGTGTATATG	ATATGACATGATGTGAGAAAGGTGTCTTGGTCCATATATGCCA	1272	
Db	1029	AGCGGTGTATATG	ATATGACATGATGTGAGAAAGGTGTCTTGGTCCATATATGCCA	1088	
QY	1273	CAGGATGGAACCA	CAATATCAATGGGTGCTTATCAAGAGAAAGTCCCTTATCCACGGCC	1332	
Db	1089	CAGGATGGAACCA	CAATATCAATGGGTGCTTATCAAGAGAAAGTCCCTTATCCACGGCC	1148	
QY	1333	AGGAATCTGTCCAG	CAAAACATTTTGGTGTGTTTGAATCTTACAAAGACCTTCTGTATGA	1392	
Db	1149	AGGAATCTGTCCAG	CAAAACATTTTGGTGTGTTTGAATCTTACAAAGACCTTCTGTATGA	1208	

QY	1393	TCGTAAACCTTTGGAGAAGAGCATCCAGCCATGTACAAATCCAGGTTTCCATATACA	1452
Db	1209	TCTTTAAACCTTTGCAGAGAGTCATCCACCATTGACANTCCAGTGTTTCTATATACAA	1268
QY	1453	TCGCCCAATAGTATCAAAACCGATGTATAATTATCAAAATTTGTCTAGACCG	1512
Db	1269	TCGCCCAATAGTATCAAAACCGATGTATAATTATCAAAATTTGTCTAGACCG	1328
QY	1513	AGTGGATCGAAGAGATGACAGATGATGTTATGTTTATCGGAACAGATGTTGGACCGT	1572
Db	1339	AGTGGATCGAAGAGATGACAGATGATGTTATGTTTATCGGAACAGATGTTGGACCGT	1388
QY	1573	TCTTAAAGTAGTTCAATTTCCTAAGAGACTGGATGATGTTTAAAGAGGTTCTGCTGGA	1632
Db	1389	TCTTAAAGTAGTTCAATTTCCTAAGAGACTGGATGATGTTTAAAGAGGTTCTGCTGGA	1448
QY	1633	AGAAATGACAGTTTTTTCGGGAACCGACTGCTATTTTCAGCAATGAGACTTTCACATPACGA	1692
Db	1449	AGAAATGACAGTTTTTTCGGGAACCGACTGCTATTTTCAGCAATGAGACTTTCACATPACGA	1508
QY	1693	GCAACACATATATATTGGTTCAACGGCTGGGGTTGCCCACTCCCTTTACACCGGTGTGA	1752
Db	1509	GCAACACATATATATTGGTTCAACGGCTGGGGTTGCCCACTCCCTTTACACCGGTGTGA	1568
QY	1753	TATTTAACGGGAAGGGGTGCTGCAATGTTTGCCTGCCCGGAGACCTTATCTGTGCTGGGA	1812
Db	1569	TATTTAACGGGAAGGGGTGCTGCAATGTTTGCCTGCCCGGAGACCTTATCTGTGCTGGGA	1628
QY	1813	TGTTTCTGCATGTTTCTCGCTATTTTCCCACTGCAAGAGACGCACAAAGCACAAGATAT	1872
Db	1629	TGTTTCTGCATGTTTCTCGCTATTTTCCCACTGCAAGAGACGCACAAAGCACAAGATAT	1688
QY	1873	AAGAAATGAGACCCCACTGACTCACTGTTTCAGACTTACACATGATTAATCAACATGSCCA	1932
Db	1689	AAGAAATGAGACCCCACTGACTCACTGTTTCAGACTTACACATGATTAATCAACATGSCCA	1748
QY	1933	CAGCCCTGAAGAGAAATCATCTAATGATGTAGAGAAATGTAGCAATTTTGGAAATGAG	1992
Db	1749	CAGCCCTGAAGAGAAATCATCTAATGATGTAGAGAAATGTAGCAATTTTGGAAATGAG	1808
QY	1993	TCCGAAGTCGCAGAGAGCGCTGCTTATTTGGCAATTCAGAGGCGCAATTAAGAGCGCAA	2052
Db	1809	TCCGAAGTCGCAGAGAGCGCTGCTTATTTGGCAATTCAGAGGCGCAATTAAGAGCGCAA	1868
QY	2053	AGAAAGATCAGAGTGAATGATCATATCATACAGACAGATCAAGGCTTCTGTACGTAG	2112
Db	1869	AGAAAGATCAGAGTGAATGATCATATCATACAGACAGATCAAGGCTTCTGTACGTAG	1928
QY	2113	TCTTCAACAGAGAAATTCAGGCAATTAACCTGTCCATCCGGTGGAAATCGGCTCATPACA	2172
Db	1929	TCTTCAACAGAGAAATTCAGGCAATTAACCTGTCCATCCGGTGGAAATCGGCTCATPACA	1988
QY	2173	AACTCTTCTTAAAGGTAAACCTTGAAGTCAATTGACACAGAGATTTGGAGAAGCTTCTTCA	2232
Db	1989	AACTCTTCTTAAAGGTAAACCTTGAAGTCAATTGACACAGAGATTTGGAGAAGCTTCTTCA	2048
QY	2233	TAAAGATGATGATGAGATGAGCTCTTAAGACCAAGAAATGTCCAATGACATGACACTAG	2292
Db	2049	TAAAGATGATGATGAGATGAGCTCTTAAGACCAAGAAATGTCCAATGACATGACACTAG	2108
QY	2293	CCAGAAAGTCTGTGTCACAGACCTTCATGACAGTCAACCAACCCCAATCTTCAACAGAT	2352
Db	2109	CCAGAAAGTCTGTGTCACAGACCTTCATGACAGTCAACCAACCCCAATCTTCAACAGAT	2168
QY	2353	GGATGAGTCTGTGTCACAGATTTGGAAAAAGGACCGAAAAACAAGTGGGCAAGGCGCAGG	2412
Db	2169	GGATGAGTCTGTGTCACAGATTTGGAAAAAGGACCGAAAAACAAGTGGGCGAAAGCCAGG	2228
QY	2413	ACATATCCCAAGGAAACAGTAAACAATGAGACACTTACAAACAAAATAGAAAGGTAGAAA	2472
Db	2229	ACATATCCCAAGGAAACAGTAAACAATGAGACACTTACAAACAAAATAGAAAGGTAGAAA	2288
QY	2473	CAGAGAGCCACAGAAATTTGAGAGGACCCAGAGAGTGTCTGAGCTGATTAACCTTACGA	2532

Db	2289	CAGAGGACCCAGCAATTTGAGAGGACCAGAGGTCTGAGCTGCAATTAACCTCTAGA	2344
Qy	2533	AACTCAAAAGTAGAAGACTTGCTAGACAAATTACTGAAAAACAAATGCAATATACAT	2592
Db	2349	AACTCAAAAGTAGAAGACTTGCTAGACAAATTACTGAAAAACAAATGCAATATACAT	2408
Qy	2533	GAACTTTTTCATGGCATTATGTGATGCTTTTACAAATGCTGGAAATTCAGCTGAGTTCCA	2655
Db	2409	GAACTTTTTCATGGCATTATGTGATGCTTTTACAAATGCTGGAAATTCAGCTGAGTTCCA	2468
Qy	2653	CCAAATTAAATTAATTAATTCATGAGTAATCTTCTTAATAGGCTTTT	2700
Db	2469	CCAAATTAAATTAATTAATTCATGAGTAATCTTCTTAATAGGCTTTT	2516
RESULT 8			
LOCUS	111900	1481 bp	DNA
DEFINITION	Sequence 1 from Patent US 5416197.	linear	PAT 26-JUL-1995
ACCESSION	111900		
VERSION	111900.1	GI:909343	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1481)		
AUTHORS	Raper, J.A. and Luo, Y.		
TITLE	Antibodies which bind human collapsin		
JOURNAL	Patent: US 5416197-A 1 16-MAY-1995;		
FEATURES	Location/Qualifiers		
source	1..1481		
BASE COUNT	454 a 299 c 345 g 383 t		
ORIGIN	/organism="unknown"		
Query Match	46.3%; Score 1253; DB 6; Length 1481;		
Best Local Similarity	99.9%; Pred. No. 0;		
Matches 1303; Conservative	0; Mismatches 1; Indels 0; Gaps 0;		
Qy	904	TGACATCTCTAAGATGACAAAGTATACCTTTCTCCGTGAAAAATGCAATAGATGAGA	963
Db	58	TGACATCTCTAAGATGACAAAGTATACCTTTCTCCGTGAAAAATGCAATAGATGAGA	117
Qy	964	ACACTCTGGAAGACTACACGCTAGATAGTCAATATGCAAGATGACCTTGGAGG	1023
Db	118	ACACTCTGGAAGAGTACTACGCTAGATAGTCAATATGCAAGATGACCTTGGAGG	177
Qy	1024	GCAACGAAGTCTGTGTAATTAATGACAACTTCTCAAAGCTGTCTGATTTGCTCGT	1088
Db	178	GCAACGAAGTCTGTGTAATTAATGACAACTTCTCAAAGCTGTCTGATTTGCTCGT	237
Qy	1084	GCCAGTCTCAATGGCATTGACACTCATTTTGTAGTAAGTGCAGAGATGATTCCTAATGAA	1143
Db	238	GCCAGTCTCAATGGCATTGACACTCATTTTGTAGTAAGTGCAGAGATGATTCCTAATGAA	297
Qy	1144	CTTTAAAGATCTTAAATTCAGTGTATATGAGGTGTTACGACTTCAGTAACTTTT	1203
Db	298	CTTTAAAGATCTTAAATTCAGTGTATATGAGGTGTTACGACTTCAGTAACTTTT	357
Qy	1204	CAAGGATCAGCGCTGTATATATAGCATAGTATGTAAGAGGGTTCCTTGGTCC	1263
Db	358	CAAGGATCAGCGCTGTATATATAGCATAGTATGTAAGAGGGTTCCTTGGTCC	417
Qy	1264	ATATGCCACAGGGATGGAACCACTATGATGGGTGCTTAATCAAGGAAGTCCCTTA	1322
Db	418	ATATGCCACAGGGATGGAACCACTATGATGGGTGCTTAATCAAGGAAGTCCCTTA	477
Qy	1324	TCCAGGCGCAGGAAGCTTGTCCCAAGAAACATTTGGTGTGTTTGACTCTACAAAGACCT	1383
Db	478	TCCAGGCGCAGGAAGCTTGTCCCAAGAAACATTTGGTGTGTTTGACTCTACAAAGACCT	537
Qy	1384	TCCTGATAGTGTATTAACCTTTGCAAGAGTCAATGAGCCATGTACATTCAGTGTTC	1443

TITLE Direct Submission
JOURNAL Submitted (01-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Mar 19, 1999 this sequence version replaced gi:1199987.

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@atcson.wustl.edu
----- Summary Statistics
Center project name: H_DU1060B11

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRN/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

This clone was derived from human PAC library RPCI-5, prepared by Pieter de Jong and coworkers at <http://www.chori.org> using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor. The clone may be obtained either from Genome Systems, Inc. (<http://www.genomesystems.com>) or Research Genetics, Inc. (<http://www.resgen.com>); or from Pieter de Jong. VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:

Actual start of this clone is at base position 1 of RP5-1060B11; actual end is at base position 179640 of RP5-1060B11.

FEATURES

source

1. 179640
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q11.23-q21.1"
/clone="RP5-1060B11"
/clone_lib="RPCI-5"
1. 866
/rpt_family="MER2_type"
229. 267
/rpt_family="AT_rich"
990. 1280
/rpt_family="Alu"
2419. 2718
/rpt_family="Alu"
2699. 2718
/rpt_family="(A)n"
10069. 10158
/rpt_family="Alu"
10297. 10328
/rpt_family="AT_rich"
10530. 10788
repeat_region

/rpt_family="Alu"
10787. 10811
/rpt_family="(CAA)n"
11346. 11448
/rpt_family="U2"
11862. 11925
/rpt_family="AT_rich"
12013. 12048
/rpt_family="(TG)n"
12096. 12301
/rpt_family="MIR"
12783. 13327
/rpt_family="ERV1"
13537. 13583
/note="match to EST BF667677 (NID:gil1941572)"
13955. 13979
/rpt_family="AT_rich"
14388. 14419
/rpt_family="(A)n"
14810. 14832
/rpt_family="AT_rich"
15750. 15833
/rpt_family="GA-rich"
16712. 16886
/rpt_family="MIR"
17541. 17698
/note="match to EST BF667677 (NID:gil1941572)"
18316. 18395
/rpt_family="(TATATG)n"
18395. 18415
/rpt_family="(T)n"
19588. 19885
/rpt_family="Alu"
19826. 20402
/note="match to EST AL134679 (NID:g6602866)"
20918. 21227
/rpt_family="Alu"
21198. 21230
/rpt_family="(A)n"
21224. 21825
/note="match to EST BG896972 (NID:gil307213)"
21724. 21746
/rpt_family="AT_rich"
21823. 21887
/note="match to EST BF667677 (NID:gil1941572)"
25562. 25640
/rpt_family="TC2"
25662. 25979
/rpt_family="U1"
25985. 26050
/rpt_family="TC2"
26845. 26884
/rpt_family="AT_rich"
26910. 27015
/rpt_family="(TA)n"
27573. 27873
/rpt_family="Alu"
27931. 27965
/rpt_family="AT_rich"
29914. 30219
/rpt_family="Alu"
30192. 30246
/rpt_family="(A)n"
30765. 31032
/note="similar to Homo sapiens EST BF356276 (NID:gil1315350)"
31015. 31224
/rpt_family="U1"
31246. 31545
/rpt_family="Alu"
31786. 31806
/rpt_family="AT_rich"
32009. 32310
repeat_region

REFERENCE
AUTHORS
TITLE
JOURNAL
University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 94027)
Waterston, R.H.
Direct Submission
Submitted (08-MAY-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
5 (bases 1 to 94027)
Waterston, R.H.
Direct Submission
Submitted (09-AUG-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
6 (bases 1 to 94027)
Waterston, R.
Direct Submission
Submitted (01-MAR-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
7 (bases 1 to 94027)
Waterston, R.
Direct Submission
Submitted (29-APR-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Aug 9, 2001 this sequence version replaced gi:13992782.

REFERENCE
AUTHORS
TITLE
JOURNAL
Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: saplens@wustl.wustl.edu
Summary Statistics
Center project name: H_NH0543D08

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send <mailto:legreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenoe, M., Catanese, J.J., and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.tegen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACE3.6.
location/Qualifiers
1..94027
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"
/map="7"

variation
/clone="RPI1-543D8"
/clone_id="RPCI-11"
895
/allele="C"
/allele="A"
/db_xref="dbSNP:2527038"
1735..1810
/rpt_family="CT-rich"
1811..2295
/rpt_family="Alu"
2342
variation
/allele="G"
/allele="A"
/db_xref="dbSNP:2691700"
2346..2377
/rpt_family="AT-rich"
2404..2431
/rpt_family="AT-rich"
3092..3122
/rpt_family="AT-rich"
4054
variation
/allele="C"
/allele="T"
/db_xref="dbSNP:2040877"
4517
variation
/allele="C"
/allele="T"
/db_xref="dbSNP:2040876"
4611..4905
/rpt_family="Alu"
5376
variation
/allele="G"
/allele="A"
/db_xref="dbSNP:2691699"
5571..5694
/rpt_family="L2"
5718..5864
/rpt_family="MIR"
5967
variation
/allele="C"
/allele="T"
/db_xref="dbSNP:2527037"
6318..6370
/rpt_family="ERV.L"
6381..6417
/rpt_family="(TA)n"
6480..6756
/rpt_family="(TATTA)n"
6788..6885
/rpt_family="MIR"
6912
variation
/allele="A"
/allele="T"
/db_xref="dbSNP:2527036"
6976
/allele="C"
/allele="C"
/db_xref="dbSNP:2527035"
7488..7510
/rpt_family="AT-rich"
7578
variation
/allele="G"
/allele="A"
/db_xref="dbSNP:1524880"
7746
variation
/allele="T"
/allele="C"
/db_xref="dbSNP:1524881"
8912
variation
/allele="A"
/allele="G"
/db_xref="dbSNP:2691698"
9493

```

repeat_region      /allele="G"
                   /allele="C"
                   /db_xref="dbSNP:2527034"
                   9525..9862 "L1"
variation           /rpt_family="L1"
                   /allele="C"
                   /allele="A"
                   /db_xref="dbSNP:2189847"
repeat_region      10203..10227
                   /rpt_family="AT_rich"
repeat_region      11126..11174
                   /rpt_family="A-rich"
repeat_region      11550..11855
                   /rpt_family="Alu"
repeat_region      12441..12462
                   /rpt_family="AT_rich"
repeat_region      12475..13649
                   /rpt_family="L1"
variation           13919
                   /allele="C"
                   /allele="T"
                   /db_xref="dbSNP:2040875"
repeat_region      14069..14124
                   /rpt_family="(TN)n"
repeat_region      14391..14438
                   /rpt_family="MER1_type"
variation           14460
                   /allele="G"
                   /allele="T"
                   /db_xref="dbSNP:2107738"

```

Query Match 8.1% Score 220; DB 9; Length 94027;

Best Local Similarity 100.0%; Pred. No. 4,2e-113; Index 0; Gaps 0;

Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1340 TGTCCAGCAAAACATTGGTGTGTTGACTCTACCAAGACCTTCTGATGATGTATA 1399
DB 87288 TGTCCAGCAAAACATTGGTGTGTTGACTCTACCAAGACCTTCTGATGATGTATA 87347

QY 1400 ACCTTTGCAAGAGTATCCAGCCATGATCCAGTGTCTTCTATGAACAATGCCCA 1459
DB 87348 ACCTTTGCAAGAGTATCCAGCCATGATCCAGTGTCTTCTATGAACAATGCCCA 87407

QY 1460 ATAGGATCAAAACGATGTAATTTATCAATTATCAAAATGTCGTAGCCAGTGGAT 1519
DB 87408 ATAGGATCAAAACGATGTAATTTATCAATTATCAAAATGTCGTAGCCAGTGGAT 87467

QY 1520 GCAGAGATGACAGATGATGTTATGTTATCGAAGCAG 1559
DB 87468 GCAGAGATGACAGATGATGTTATGTTATGTTATCGAAGCAG 87507

```

```

RESULT 13
AC004451/c 108642 bp DNA linear PRI 04-JUN-2002
LOCUS Homo sapiens PAC clone RP4-789N1 from 7q21, complete sequence.
DEFINITION AC004451
ACCESSION AC004451
VERSION AC004451.2 GI:21322198
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 108642)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 108642)
AUTHORS Woessner, J., Minx, P., Hinds, K. and Stromwater, C.
TITLE The sequence of Homo sapiens PAC clone RP4-789N1

```

```

JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 108642)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (20-MAR-1998) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 4 (bases 1 to 108642)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE 5 (bases 1 to 108642)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (04-JUN-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Jun 4, 2002 this sequence version replaced gi:2979604.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: MUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H_DJ0789N01
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

This clone was derived from human PAC library RPCI-4, prepared by Pieter de Jong and coworkers at <http://www.chori.org> using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor. The clone may be obtained either from Genome Systems, Inc. (<http://www.genomesystems.com>) or Research Genetics, Inc. (<http://www.resgen.com>); or from Pieter de Jong. VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP4-649P17; the clone sequenced to the right is RP4-800B9, 200 bp overlap. Actual start of this clone is at base position 78831 of RP4-649P17; actual end is at base position 108516 of RP4-789N1.

FEATURES

```

source
1..108642
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"
/map="7q21"
/clone="RP4-789N1"
/clone_lib="RPCI-4"

```

```

repeat_region      1366..1394      /rpt_family="AT_rich"
repeat_region      3626..3692      /rpt_family="(GA)n"
repeat_region      4049..4087      /rpt_family="(CA)n"
repeat_region      4975..5033      /rpt_family="(GA)n"
repeat_region      5063..5123      /rpt_family="purine-rich"
repeat_region      5481..5519      /rpt_family="AT_rich"
repeat_region      6368..6676      /rpt_family="Alu"
repeat_region      6820..7138      /rpt_family="Alu"
repeat_region      10797..10996      /rpt_family="L1"
repeat_region      11925..12010      /rpt_family="L2"
repeat_region      12140..12164      /rpt_family="(TA)n"
repeat_region      12165..12467      /rpt_family="Alu"
repeat_region      12793..12823      /rpt_family="AT_rich"
repeat_region      14958..15258      /rpt_family="Alu"
repeat_region      15334..15386      /rpt_family="Alu"
repeat_region      16602..16720      /rpt_family="(TA)n"
repeat_region      16808..17098      /rpt_family="Alu"
repeat_region      17939..17978      /rpt_family="(CA)n"
repeat_region      18063..18146      /rpt_family="MIR"
repeat_region      18419..18453      /rpt_family="AT_rich"
repeat_region      18802..19345      /rpt_family="L2"
repeat_region      19352..19382      /rpt_family="POLY_A"
repeat_region      19366..19727      /note="match to EST AA731945 (NID:g2753896) nm6490.s1"
repeat_region      19367..19692      /note="similar to EST AA644250 (NID:g2569468) af63b06.s1"
repeat_region      19368..19727      /note="similar to EST AA79841 (NID:g2839172) af45g01.s1"
repeat_region      19378..19791      /note="similar to EST A1122887 (NID:g3538653) qb01f04.s1"
repeat_region      19577..19791      /note="match to EST AA320720 (NID:g1973048) "
repeat_region      19611..19791      /note="similar to EST AA171850 (NID:g1751104) z098f08.s1"
repeat_region      19966..20271      /rpt_family="Alu"
repeat_region      21538..21564      /rpt_family="AT_rich"
repeat_region      21605..21913      /rpt_family="Alu"
repeat_region      21924..22322      /rpt_family="Alu"
repeat_region      22714..22884      /rpt_family="MIR"
repeat_region      23335..23636      /rpt_family="Alu"
repeat_region      23637..23807      /rpt_family="Alu"
repeat_region      23813..23835      /rpt_family="(TAAA)n"
repeat_region      23845..24126

```

```

repeat_region      24156..24398      /rpt_family="Alu"
repeat_region      24717..25041      /rpt_family="L2"
repeat_region      26317..26414      /rpt_family="Alu"
repeat_region      26433..26627      /rpt_family="Alu"
repeat_region      26703..26729      /rpt_family="Alu"
repeat_region      26894..26947      /rpt_family="AT_rich"
repeat_region      27824..27859      /rpt_family="AT_rich"
repeat_region      28102..28368      /rpt_family="AT_rich"
repeat_region      28341..28500      /rpt_family="MER2_type"
repeat_region      28500..30640      /rpt_family="MER2_type"
repeat_region      31284..31318      /rpt_family="AT_rich"
repeat_region      31447..31584      /rpt_family="MER1_type"
repeat_region      31928..32008      /rpt_family="(CAT)n"
repeat_region      35757..35850      /rpt_family="(TA)n"
repeat_region      36529..36559      /rpt_family="(CA)n"
repeat_region      37257..37640      /rpt_family="retroviral"
repeat_region      38296..38328      /rpt_family="AT_rich"
repeat_region      38501..39044

```

Query Match 7.2%; Score 194; DB 9; Length 108642;
 Best Local Similarity 100.0%; Pred. No. 2,9e-98;
 Matches 194; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 118 TTTCAGATTGTTGAACTTCTCTGCGCAGCAATACAGAGAGACATAAGCAGCAA 177
DB 3305 TTTCAGATTGTTGAACTTCTCTGCGCAGCAATACAGAGAGACATAAGCAGCAA 3246
QY 178 AGGAACTTACAGCCTTCGACGATGGCGTGTACTAGATTTGTCTTTCTGGGG 237
DB 3245 AGGAACTTACAGCCTTCGACGATGGCGTGTACTAGATTTGTCTTTCTGGGG 3186
QY 238 AGTATTAATTACAGAGAGAGAACTATCAGAAATGGAGAGAAATGCGCAAGCTGAA 297
DB 3185 AGTATTAATTACAGAGAGAGAACTATCAGAAATGGAGAGAAATGCGCAAGCTGAA 3126
QY 298 ATTATCTTACAAAG 311
DB 3125 ATTATCTTACAAAG 3112

```

RESULT 14
 AC004848/c 135214 bp DNA linear PRI 26-APR-2003
 LOCUS Homo sapiens PAC clone RP4-649P17 from 7, complete sequence.
 DEFINITION AC004848
 ACCESSION AC004848
 VERSION AC004848.1 GI:3900858
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 135214)
 AUTHORS Sulston, J.E. and Waterston, R.
 TITLE Toward a complete human genome sequence
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)

MEDLINE 99063792
 PUBMED 9847074
 REFERENCE 2 (bases 1 to 135214)
 AUTHORS Duclels,G., Rohlfing,T. and Tin-wollam,A.M.
 TITLE The sequence of Homo sapiens PAC clone RP4-649P17
 JOURNAL Unpublished (2001)
 REFERENCE 3 (bases 1 to 135214)
 AUTHORS Waterston,R.H.
 TITLE Direct Submission
 JOURNAL Submitted (12-JUN-1998) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 4 (bases 1 to 135214)
 AUTHORS Waterston,R.
 TITLE Direct Submission
 JOURNAL Submitted (21-NOV-1998) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 REFERENCE 5 (bases 1 to 135214)
 AUTHORS Waterston,R.
 TITLE Direct Submission
 JOURNAL Submitted (26-APR-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 REFERENCE 6 (bases 1 to 135214)
 AUTHORS Waterston,R.
 TITLE Direct Submission
 JOURNAL Submitted (26-APR-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
 COMMENT On Nov 21, 1998 this sequence version replaced gi:3213149.
 ----- Genome Center
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site: http://genome.wustl.edu/gsc
 Contact: saplens@wustl.wustl.edu
 ----- Summary Statistics
 Center project name: H_DJ0649P17

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
 The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see
 http://www.nhgri.nih.gov/DIR/GRB/CHR7, send
 mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
 This clone was derived from human PAC library RPCI-4, prepared by Pieter de Jong and coworkers at http://www.chori.org using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor.
 The clone may be obtained either from Genome Systems, Inc. (http://www.genomesystems.com) or Research Genetics, Inc. (http://www.resgen.com), or from Pieter de Jong.
 VECTOR: pCYPAC2

NEIGHBORING SEQUENCE INFORMATION:
 The clone sequenced to the right is RP4-789N1, 200 bp overlap.
 Actual start of this clone is at base position 1 of RP4-649P17

FEATURES actual end is at 135018 of DJ0649P17.
 source Location/Qualifiers
 1..135214

/organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="7"
 /map="7"

/clone="RP4-649P17"
 /clone_1db="RPCI-4"

repeat_region 261..298
 /rpt_family="(CA)n"

variation 483
 /allele="A"
 /allele="G"
 /db_xref="dbSNP:2527040"

repeat_region 598..691
 /rpt_family="MIR"
 /rpt_family="MIR"

variation 1021
 /allele="T"
 /allele="C"
 /db_xref="dbSNP:2527030"

repeat_region 1136..1402
 /rpt_family="Alu"
 1403..1443
 /rpt_family="(TAA)n"

variation 1957
 /allele="C"
 /allele="T"

repeat_region 3187..3240
 /rpt_family="AT-rich"
 3415..3465
 /rpt_family="A-rich"

variation 3525
 /allele="G"
 /allele="A"

variation 3674
 /db_xref="dbSNP:2189848"

variation 3941
 /allele="T"
 /allele="A"

variation 3949
 /db_xref="dbSNP:2301890"

variation 3949
 /allele="C"
 /allele="T"

variation 3949
 /db_xref="dbSNP:2527039"

variation 3949
 /db_xref="dbSNP:2527039"

variation 3949
 /db_xref="dbSNP:2527039"

variation 3949
 /db_xref="dbSNP:2527039"

variation 3949
 /db_xref="dbSNP:2527039"

variation 3949
 /db_xref="dbSNP:2527039"

variation 3949
 /db_xref="dbSNP:2527039"

variation 3949
 /db_xref="dbSNP:2527039"

variation 3949
 /db_xref="dbSNP:2527039"

variation 3949
 /db_xref="dbSNP:2527039"

variation 3949
 /db_xref="dbSNP:2527039"

```

variation      /db_xref="dbSNP:1157309"
5036
/allele="A"
/allele="C"
/db_xref="dbSNP:2691702"
variation      5078
/allele="A"
/allele="G"
/db_xref="dbSNP:2527041"
repeat_region  5723..5751
/rpt_family="AT_rich"
variation      5906
/allele="C"
/allele="A"
/db_xref="dbSNP:2189849"
repeat_region  5953..5973
/rpt_family="AT_rich"
variation      5978
/allele="A"
/allele="G"
/db_xref="dbSNP:2189850"
6078..6504
/nc="CpG_island (%GC=59.3, o/e=0.89, #CpGs=33)"
misc_feature   6095
/allele="A"
/allele="G"
/db_xref="dbSNP:2691701"
variation      6127
/allele="G"
/allele="C"
/db_xref="dbSNP:2527042"
variation      12512
/allele="T"
/allele="C"
/db_xref="dbSNP:2527031"
12851..12970
/rpt_family="MIR"
variation      14674
/allele="C"
/allele="T"
/db_xref="dbSNP:2888220"

Query Match      5.9%; Score 160; DB 9; Length 135214;
Best Local Similarity 100.0%; Pred. No. 7.4e-79;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGATCCAAACATGTGATCACTTTCATGCTGGCCAAACAGCTCCAGTTA 369
|||||
Db 78604 AGAATGTTGGATCCAAACATGTGATCACTTTCATGCTGGCCAAACAGCTCCAGTTA 78545

QY 370 TCATACCTTCCTTTGGATGAGAGAGAGTAGGCTGTATGTTGAGAGAGATCACAT 429
|||||
Db 78544 TCATACCTTCCTTTGGATGAGAGAGAGTAGGCTGTATGTTGAGAGAGATCACAT 78485

QY 430 ATTTTCATTCGACCTGTTAATATCAAGATTTTCAAAAG 469
|||||
Db 78484 ATTTTCATTCGACCTGTTAATATCAAGATTTTCAAAAG 78445

RESULT 15
AC073110 169408 bp DNA linear HTG 03-FEB-2001
LOCUS AC073110 Homo sapiens chromosome 7 clone RP11-506C5, WORKING DRAFT SQUIDENCE,
DEFINITION 21 unordered pieces.
ACCESSION AC073110
VERSION AC073110.5 GI:12658189
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 169408)
AUTHORS Waterston,R.H.

```

```

TITLE      The sequence of Homo sapiens clone
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 169408)
AUTHORS    Waterston,R.H.
TITLE      Direct Submission
JOURNAL    Submitted (08-JUN-2000) Genome Sequencing Center, Washington
COMMENT    University School of Medicine, 4444 Forest Park Parkway, St. Louis,
           MO 63108, USA
           On Feb 3, 2001 this sequence version replaced gi:11192243.

-----Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
-----Project Information -----
Center project name: H.NH0506C05
-----Summary Statistics -----
Sequencing vector: M13; 84%
Sequencing vector: plasmid; 5%
Chemistry: Dye-primer ET; 84% of reads
Chemistry: Dye-terminator Big Dye; 5% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 157223 bases at least Q40
Consensus quality: 160650 bases at least Q30
Consensus quality: 162652 bases at least Q20
Insert size: 155000; agarose-fp
Insert size: 166528; sum-of-contigs
Quality coverage: 4.21 in Q20 bases; agarose-fp
Quality coverage: 4.09 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1260: contig of 1260 bp in length
* 1261 1360: gap of unknown length
* 1361 2625: contig of 1265 bp in length
* 2626 2725: gap of unknown length
* 2726 4145: contig of 1420 bp in length
* 4146 4246: gap of unknown length
* 4246 5681: contig of 1436 bp in length
* 5682 5781: gap of unknown length
* 5782 7286: contig of 1505 bp in length
* 7287 7386: gap of unknown length
* 7387 10841: contig of 3455 bp in length
* 10842 10941: gap of unknown length
* 10942 15222: contig of 4281 bp in length
* 15223 15322: gap of unknown length
* 15323 16202: contig of 880 bp in length
* 16203 16303: gap of unknown length
* 16303 19851: contig of 3549 bp in length
* 19852 19951: gap of unknown length
* 19952 24367: contig of 4416 bp in length
* 24368 24468: gap of unknown length
* 24469 30317: contig of 5850 bp in length
* 30318 30418: gap of unknown length
* 30419 36826: contig of 6409 bp in length
* 36827 36926: gap of unknown length
* 36927 43388: contig of 6463 bp in length
* 43389 43489: gap of unknown length
* 43490 50499: contig of 7010 bp in length
* 50500 50599: gap of unknown length
* 50600 57156: contig of 6557 bp in length
* 57157 57256: gap of unknown length
* 57257 65679: contig of 8423 bp in length
* 65680 65779: gap of unknown length
* 65780 76066: contig of 10287 bp in length
* 76067 76167: gap of unknown length
* 76168 89540: contig of 13374 bp in length

```

FEATURES

89541 89640: gap of unknown length

* 89641 110939: contig of 21299 bp in length

* 110940 111039: gap of unknown length

* 111040 135938: contig of 24899 bp in length

* 135939 136038: gap of unknown length

* 136039 169408: contig of 33370 bp in length.

Location/Qualifiers

source

1. 169408

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="7"

/clone="RP11-506C5"

1. 1260

/note="assembly_name:Contig11"

1361. 2625

/note="assembly_name:Contig23"

2726. 4145

/note="assembly_name:Contig34"

4246. 5681

/note="assembly_name:Contig35"

5782. 7286

/note="assembly_name:Contig37"

7387. 10841

/note="assembly_name:Contig40"

10942. 15222

/note="assembly_name:Contig41

clone_end:SP6

vector_side:right"

15323. 16202

/note="assembly_name:Contig3"

16303. 19851

/note="assembly_name:Contig42"

19952. 24367

/note="assembly_name:Contig43"

24468. 30317

/note="assembly_name:Contig44"

30418. 36826

/note="assembly_name:Contig45"

36927. 43389

/note="assembly_name:Contig46"

43490. 50499

/note="assembly_name:Contig47"

50600. 57156

/note="assembly_name:Contig48"

57257. 65679

/note="assembly_name:Contig49"

65780. 76066

/note="assembly_name:Contig50"

76167. 89540

/note="assembly_name:Contig51"

89641. 110939

/note="assembly_name:Contig52"

111040. 135938

/note="assembly_name:Contig53"

136039. 169408

/note="assembly_name:Contig54"

BASE COUNT 52692 a 29667 c 29411 g 55592 t 2046 others

ORIGIN

Query Match 5.5%; Score 148; DB 2; Length 169408;

Best Local Similarity 100.0%; Pred. No. 5; 2e-72;

Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1132 CAGTACATTTCAGAGGATGCGCTGTGTATGATACAGAGTGTAGAGGCT 1251

DB 49274 CAGTACATTTCAGAGGATGCGCTGTGTATGATACAGAGTGTAGAGGCT 49333

QY 1252 GTTCTGTGTCATATGCGCAGAGGATGAGCCCACTATCAATGAGTCTTATCAAG 1311

DB 49334 GTTCTGTGTCATATGCGCAGAGGATGAGCCCACTATCAATGAGTCTTATCAAG 49393

QY 1312 AAGAGTCCCTATCCAGGCGCAGGAAGT 1339

Db 49394 AAGAGTCCCTATCCAGGCGCAGGAAGT 49421

RESULT 16

AC091270/c

LOCUS

DEFINITION

AC091270 Homo sapiens chromosome 11 clone RP11-115M2 map 11, LOW-PASS

SEQUENCE SAMPLING.

AC091270.1 GI:13605997

AC091270

HTG: HTGS_PHASE0.

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE

AUTHORS

1 (bases 1 to 81667)

Birren, B., Linton, L., Nusbaum, C. and Lander, E.

REFERENCE

TITLE

JOURNAL

REFERENCE

AUTHORS

2 (bases 1 to 81667)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,

Barna, N., Baethen, V., Boguslavsky, L., Boukhalter, B., Brown, A.,

Camarda, D., Campiano, A., Chang, J., Choquet, J., Colangelo, M.,

Collins, S., Collymore, A., Cooke, P., Deaellano, K., Dewar, K.,

Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D.,

Galagan, J., Gardyna, S., Glade, S., Goyette, M., Graham, L.,

Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W.,

Iliev, I., Johnson, R., Jones, C., Karatas, A., Labocque, K.,

Lamarez, R., Landers, T., Lehotzky, J., Levine, R., Liu, G.,

Maclean, C., Macdonald, P., Margulis, N., Matthews, C., McCarthy, M.,

McEwan, P., McKernan, K., McPheters, R., Melitini, J., Menus, L.,

Mihova, T., Mleng, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C.,

Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,

Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,

Retica, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,

Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S.,

Severy, P., Sougnuez, C., Spencer, B., Stange-Thomann, N.,

Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S.,

Theodore, J., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,

Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,

Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL

COMMENT

Submitted (12-Apr-2001) Whitehead Institute/MIT Center for Genome

Research, 320 Charles Street, Cambridge, MA 02141, USA

All repeats were identified using RepeatMasker:

Smt, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L13092

Center clone name: L15_M_2

NOTE: This record contains 70 individual

* sequencing reads that have not been assembled into

* contigs. Runs of N are used to separate the reads

* and the order in which they appear is completely

* arbitrary. Low-pass sequence sampling is useful for

* identifying clones that may be gene-rich and allows

* overlap relationships among clones to be deduced.

* However, it should not be assumed that this clone

* will be sequenced to completion. In the event that

* the record is updated, the accession number will

* be preserved.

1 1061: contig of 1061 bp in length

* 1062 1161: gap of 100 bp

* 1162 2202: contig of 1041 bp in length

* 2203 2302: gap of 100 bp

* 2303 3358: contig of 1056 bp in length

3359 3458: gap of 100 bp
3459 4520: contig of 1062 bp in length
4521 4620: gap of 100 bp
4621 5725: contig of 1105 bp in length
5726 5825: gap of 100 bp
5826 6895: contig of 1070 bp in length
6896 6995: gap of 100 bp
6996 8043: contig of 1048 bp in length
8044 8143: gap of 100 bp
8144 9212: contig of 1069 bp in length
9213 9312: gap of 100 bp
9313 10439: contig of 1127 bp in length
10440 10539: gap of 100 bp
10540 11612: contig of 1073 bp in length
11613 11712: gap of 100 bp
11713 12750: contig of 1038 bp in length
12751 12850: gap of 100 bp
12851 13886: contig of 1036 bp in length
13887 13986: gap of 100 bp
13987 15027: contig of 1041 bp in length
15028 15127: gap of 100 bp
15128 16197: contig of 1070 bp in length
16198 16297: gap of 100 bp
16298 17398: contig of 1101 bp in length
17399 17498: gap of 100 bp
17499 18610: contig of 1112 bp in length
18611 18710: gap of 100 bp
18711 19754: contig of 1044 bp in length
19755 19855: gap of 100 bp
19855 20944: contig of 1090 bp in length
20945 21044: gap of 100 bp
21045 22045: contig of 1000 bp in length
22046 22144: gap of 100 bp
22145 23221: contig of 1077 bp in length
23222 23321: gap of 100 bp
23322 24304: contig of 1083 bp in length
24305 25404: gap of 100 bp
25405 25643: contig of 1139 bp in length
25644 25743: gap of 100 bp
25744 26660: contig of 1117 bp in length
26661 26960: gap of 100 bp
26961 28028: contig of 1068 bp in length
28029 28128: gap of 100 bp
28129 29216: contig of 1088 bp in length
29217 29316: gap of 100 bp
29317 30393: contig of 1077 bp in length
30394 30493: gap of 100 bp
30494 31584: contig of 1091 bp in length
31585 32761: contig of 1077 bp in length
32762 32861: gap of 100 bp
32863 33951: contig of 1090 bp in length
33952 34051: gap of 100 bp
34052 35075: contig of 1024 bp in length
35076 35175: gap of 100 bp
35176 36272: contig of 1097 bp in length
36273 36372: gap of 100 bp
36373 37441: contig of 1069 bp in length
37442 37541: gap of 100 bp
37542 38600: contig of 1059 bp in length
38601 38700: gap of 100 bp
38701 39781: contig of 1081 bp in length
39782 40979: contig of 1098 bp in length
40980 41079: gap of 100 bp
41080 42096: contig of 1017 bp in length
42097 42196: gap of 100 bp
42197 43295: contig of 1099 bp in length
43296 43395: gap of 100 bp
43396 44461: contig of 1066 bp in length
44462 44561: gap of 100 bp
44562 45625: contig of 1064 bp in length
45626 45725: gap of 100 bp

45726 46830: contig of 1105 bp in length
46831 46930: gap of 100 bp
46931 48033: contig of 1103 bp in length
48034 48133: gap of 100 bp
48134 49181: contig of 1048 bp in length
49182 49281: gap of 100 bp
49282 50381: contig of 1100 bp in length
50382 50481: gap of 100 bp
50482 51600: contig of 1119 bp in length
51601 51700: gap of 100 bp
51701 52817: contig of 1117 bp in length
52818 52917: gap of 100 bp
52918 53967: contig of 1050 bp in length
53968 54067: gap of 100 bp
54068 55103: contig of 1036 bp in length
55104 55203: gap of 100 bp
55204 56264: contig of 1061 bp in length
56265 56364: gap of 100 bp
56365 57469: contig of 1105 bp in length
57470 57569: gap of 100 bp
57570 58636: contig of 1067 bp in length
58637 58736: gap of 100 bp
58737 59797: contig of 1061 bp in length
59798 59897: gap of 100 bp
59898 60979: contig of 1082 bp in length
60980 61079: gap of 100 bp
61080 62035: contig of 956 bp in length
62036 62135: gap of 100 bp
62136 63164: contig of 1029 bp in length
63165 63264: gap of 100 bp
63265 64305: contig of 1041 bp in length
64306 64405: gap of 100 bp
64406 65427: contig of 1022 bp in length
65428 65527: gap of 100 bp
65528 66574: contig of 1047 bp in length
66575 66674: gap of 100 bp
66676 67683: contig of 1009 bp in length
67684 67783: gap of 100 bp
67784 68875: contig of 1093 bp in length
68876 68975: gap of 100 bp
68977 69998: contig of 1022 bp in length
69999 70098: gap of 100 bp
70099 71149: contig of 1051 bp in length
71150 71249: gap of 100 bp
71250 72339: contig of 1090 bp in length
72340 72439: gap of 100 bp
72440 73511: contig of 1072 bp in length
73512 74726: contig of 1115 bp in length
74727 74826: gap of 100 bp
74827 75864: contig of 1038 bp in length
75865 75964: gap of 100 bp
75965 77015: contig of 1051 bp in length
77016 77115: gap of 100 bp
77116 78183: contig of 1068 bp in length
78184 78283: gap of 100 bp
78284 79298: contig of 1015 bp in length
79299 79398: gap of 100 bp
79399 80451: contig of 1053 bp in length
80452 80551: gap of 100 bp

Query Match 4.3%; Score 117; DB 2; Length 81667;
Best Local Similarity 100.0%; Pred. No. 2,7e-54; Indels 0; Gaps 0;
Matches 117; Conservative 0; Mismatches 0;

Qy 1008 AGAATGACTTTGAGGGCAGAGAGTGTGTGTAATAATGACACATTCTCAAGCTC 1067
Db 20163 AGAATGACTTTGAGGGCAGAGAGTGTGTGTAATAATGACACATTCTCAAGCTC 20104
Qy 1068 GTCTGATTTCTCAGTCCAGTCCAAATGCGATTGACATCTCTTTGATGAAGTGC 1124
Db 20103 GTCTGATTTCTCAGTCCAGTCCAAATGCGATTGACATCTCTTTGATGAAGTGC 20047

```

RESULT 17
AX207159      2137 bp      DNA      linear      PAT 30-AUG-2001
LOCUS         Sequence 6 from Patent WO0155455.
DEFINITION    AX207159
VERSION       AX207159.1 GI:15394947
KEYWORDS
SOURCE        Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS       Jin, S.
TITLE         Resistance sequences and uses thereof.
JOURNAL       Patent: WO 0155455-A 6 02-AUG-2001;
              Millennium Pharmaceuticals, Inc. (US) ; Jin, Shengfang (US)
FEATURES
source        1..2137
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
BASE COUNT    658 a 454 c 428 g 597 t
ORIGIN
Query Match   2.6%; Score 71; DB 6; Length 2137;
Best Local Similarity 100.0%; Pred. No. 6, 1e-28;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATGATGTTAACAAGCTTAATTCAGTCGACGCGGCGTGC 60
DB 17 AATCTTTATTTATGATGTTAACAAGCTTAATTCAGTCGACGCGGCGTGC 76
QY 61 ACCCAGCGCTC 71
DB 77 ACCCAGCGCTC 87

RESULT 18
RNSIIICM1    2319 bp      mRNA      linear      ROD 29-JAN-1996
LOCUS         R.norvegicus mRNA for semaphorin III/collapsin-1.
DEFINITION    X95286
VERSION       X95286.1 GI:1171618
KEYWORDS     collapsin-1; semaphorin III.
SOURCE       Rattus norvegicus (Norway rat)
ORGANISM     Rattus norvegicus
              Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
              Rattus.

REFERENCE
AUTHORS       Giger, R.I.
JOURNAL       Unpublished
REFERENCE     2 (bases 1 to 2319)
AUTHORS       Giger, R.I.
TITLE         Direct Submision
JOURNAL       Submitted (22-JAN-1996) R.I. Giger, Netherlands Inst. for Brain
              Research, Weibergdreef 33, 1105 AZ, Amsterdam, NETHERLANDS
FEATURES
source        1..2319
              /organism="Rattus norvegicus"
              /mol_type="mRNA"
              /strain="Wistar"
              /db_xref="taxon:10116"
              /tissue_type="brain"
              /dev_stage="adult"
              1..2319
              /codon_start=1
              /product="semaphorin III/collapsin-1"
              /protein_id="CAA64607.1"
              /db_xref="GI:1171618"
              /db_xref="SPTREMBL:O63548"
              /translation="MGWFTGIACLFWGIILTARAVYANGKNVPRILKSYKMELESNN"

CDS

```

```

VITFNGLANSSVHTFLDEERSRLVYAKDHIESFNLWNIKDFQKIVWPVSYTRDE
CKMAKGDILKCCANFIKYLKAYNOTHLAYACGAPHPICITLYEVGHHEEDNIFKQDS
HFNENGRKSPYDPKLITASLIDGLVSGTADPEGRPAIFRTLGHHEDNIFKQDS
RVLNDPRISAHILIDSDNPPDDKYFPERFENALIDGHSRGAHTARIGOLICNDGCH
KSLVKNWTFILKARLIGCPGPNGLIDHFDLIDVFLNNSDKPRPIYVGFPTSSNI
RSGAVCMYSMSDVARFVGLPAHRDGNVQGVVPRPQTCSTKTFEGSDST
KDLPDVITFARSHPAMTNPVPEFINNRPIMKITDVNVOFTQIVDRAVEDGQDVMF
IGTDVGLTVKSVKPMETMHDLEVLLEMTVFEPTTISAMELSTKQOQYIGSTAG
VAQLPILHRCDIYGKACAECCCLADPYCAMDSSCSRYPPTAKRRTRRODIRNGDLTH
CSDLOHNDHNGHSLERILYGVENSSTFLCSPSOPALVYMOPORNRNDRKERTIV
GDHTRTQGLILSLRLOKDSGNVICHAVEHGFNOTLKVTLVETDTHLESLHKDD
DDGDSKTKEMSSSMTPSQVWYRDMQILNPNLNTMDEFCBQVWKRURKORRQPH
SQSSNKKWKMHOESKGRNRTTHERPERSV"

BASE COUNT    704 a 491 c 548 g 576 t
ORIGIN
Query Match   2.5%; Score 68; DB 10; Length 2319;
Best Local Similarity 100.0%; Pred. No. 3, 1e-26;
Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 TACAAGAAATGTTGAATCCACAATGATGATTCATTCAATGCGTGGCCACAGCTCC 364
DB 106 TACAAGAAATGTTGAATCCACAATGATGATTCATTCAATGCGTGGCCACAGCTCC 165
QY 365 AGTTATCA 372
DB 166 AGTTATCA 173

RESULT 19
BD026302     229 bp      DNA      linear      PAT 27-AUG-2002
LOCUS         BD026302
DEFINITION    Sequence tag and encoded human protein.
ACCESSION    BD026302
VERSION      BD026302.1 GI:22567525
KEYWORDS     JP 2001269182-A/2548.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS       Edwards, J.B.D.M., Duclair, E. and Jordan, J.V.
JOURNAL       Sequence tag and encoded human protein
              Patent: JP 2001269182-A 2548 02-OCT-2001;
              GENSET
COMMENT
OS Homo sapiens (human)
PN JP 2001269182-A/2548
PD 02-OCT-2001
PF 24-FEB-2000 JP 2000118773
PR 26-FEB-1999 US 60/122487
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, EIMERIC DUCLAIR, JEAN YVES
PI JORDAN
PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC
C12N5/10,
PC C12P21/02, C12P21/08, C12Q1/68//G06F17/30, C12N15/00, C12N5/00, PC
G06F15/40
CC Key Location/Qualifiers
FH 29..229.
FT CDS
FEATURES
source        1..229
              /organism="Homo sapiens"
              /mol_type="genomic DNA"
              /db_xref="taxon:9606"
BASE COUNT    63 a 47 c 53 g 62 t 4 others
ORIGIN
Query Match   2.3%; Score 63; DB 6; Length 229;
Best Local Similarity 100.0%; Pred. No. 2, 5e-23;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 174 GCAAGGACCTACAGCGTTCGACAGATGGCGTGTAACTAGATGCTGCTTTCT 233

```

Db 123 GCAAGGACCTACAGCGCTGCAGCAGCGCTGTTAACTAGATTGTCGCTTTCT 182
Qy 234 GGG 236
Db 183 GGG 185

RESULT 20

AC109862.2/c
WPCOMMENT

Sequence split into 4 fragments LOCUS AC109862 Accession AC109862

Fragment Name Begin End
AC109862_0 1 110000
AC109862_1 100001 210000
AC109862_2 200001 310000
AC109862_3 300001 381240
Continuation (3 of 4) of AC109862 from base 200001 (AC109862 Rattus norvegicus clone CH2)

Query Match 2.3%; Score 63; DB 2; Length 110000;
Best Local Similarity 100.0%; Pred. No. 1.8e-23;
Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 310 AGAATCTTGAATCCAGATGATCATCTTCAATGCTTGGCCACAGCTCCAGTTA 369
Db 20115 AGAATGTTGAAATCCAAATGATGATCATCTTCAATGCTTGGCCACAGCTCCAGTTA 20056
Qy 370 TCA 372
Db 20055 TCA 20053

RESULT 21
LOCUS AC091270 81667 bp DNA linear HTG 12-APR-2001
DEFINITION Homo sapiens chromosome 11 clone RP11-115M2 map 11, LOW-PASS
SEQUENCE SAMPLING.
AC091270
AC091270.1 GI:13605997
VERSION HTG; HTGS PHASE0.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
JOURNAL 1 (bases 1 to 81667)
REFERENCE Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens chromosome 11, clone RP11-115M2
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 81667)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Barna, N., Baetien, V., Boguslavskiy, L., Bouckgalter, B., Brown, A.,
Camata, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M.,
Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K.,
Diaz, T. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Hulme, W.,
Iliev, I., Johnson, R., Jones, C., Karatas, A., Labroque, K.,
Lamarez, R., Landers, T., Lehotzky, J., Levine, R., Liu, G.,
Maclean, C., Macdonald, P., Marguis, N., Matthews, C., McCarthy, M.,
McGowan, P., McKernan, K., McPheters, R., Melidim, J., Menus, L.,
Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norby, C.,
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
Retta, R., Ribback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnapack, R., Seaman, S.,
Sever, P., Sougniez, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Strauss, N., Subramanian, A., Talama, J., Testaye, S.,
Theodore, J., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
JOURNAL Submitted (12-APR-2001) Whitehead Institute/MIT Center for Genome
COMMENT Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIGR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L13092
Center clone name: 115_M_2

* NOTE: This record contains 70 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 1061: contig of 1061 bp in length
1062 1161: gap of 100 bp
1162 2202: contig of 1041 bp in length
2203 2302: gap of 100 bp
2303 3358: contig of 1056 bp in length
3359 3458: gap of 100 bp
3459 4520: contig of 1062 bp in length
4521 4620: gap of 100 bp
4621 5725: contig of 1105 bp in length
5726 5825: gap of 100 bp
5826 6895: contig of 1070 bp in length
6896 6995: gap of 100 bp
6996 8043: contig of 1048 bp in length
8044 8143: gap of 100 bp
8144 9212: contig of 1069 bp in length
9213 9312: gap of 100 bp
9313 10439: contig of 1127 bp in length
10440 10539: gap of 100 bp
10540 11612: contig of 1073 bp in length
11613 11712: gap of 100 bp
11713 12750: contig of 1038 bp in length
12751 12850: gap of 100 bp
12851 13865: contig of 1036 bp in length
13867 13986: gap of 100 bp
13987 15027: contig of 1041 bp in length
15028 15127: gap of 100 bp
15128 16197: contig of 1070 bp in length
16198 16297: gap of 100 bp
16299 17398: contig of 1101 bp in length
17399 17498: gap of 100 bp
17499 18610: contig of 1112 bp in length
18611 18710: gap of 100 bp
18711 19754: contig of 1044 bp in length
19755 19854: gap of 100 bp
19855 20944: contig of 1090 bp in length
20945 21044: gap of 100 bp
21045 22044: contig of 1000 bp in length
22045 22144: gap of 100 bp
22145 23221: contig of 1077 bp in length
23222 23321: gap of 100 bp
23322 24404: contig of 1083 bp in length
24405 24504: gap of 100 bp
24505 25643: contig of 1139 bp in length
25644 25743: gap of 100 bp
25744 26860: contig of 1117 bp in length
26861 26960: gap of 100 bp
26961 28028: contig of 1068 bp in length
28029 28128: gap of 100 bp
28129 29216: contig of 1088 bp in length
29217 29316: gap of 100 bp
29317 30393: contig of 1077 bp in length

```

* 30394 30493: gap of 100 bp
* 30494 31584: contig of 1091 bp in length
* 31584 31684: gap of 100 bp
* 31684 32761: contig of 1077 bp in length
* 32761 32861: gap of 100 bp
* 32861 33951: contig of 1090 bp in length
* 33951 34051: gap of 100 bp
* 34051 35076: contig of 1024 bp in length
* 35076 35176: gap of 100 bp
* 35176 36272: contig of 1097 bp in length
* 36272 36372: gap of 100 bp
* 36372 37441: contig of 1069 bp in length
* 37441 37541: gap of 100 bp
* 37541 38600: contig of 1059 bp in length
* 38600 38700: gap of 100 bp
* 38700 39781: contig of 1081 bp in length
* 39781 39881: gap of 100 bp
* 39881 40979: contig of 1098 bp in length
* 40979 41079: gap of 100 bp
* 41079 42096: contig of 1017 bp in length
* 42096 42196: gap of 100 bp
* 42196 43295: contig of 1099 bp in length
* 43295 43395: gap of 100 bp
* 43395 44461: contig of 1066 bp in length
* 44461 44561: gap of 100 bp
* 44561 45625: contig of 1064 bp in length
* 45625 45725: gap of 100 bp
* 45725 46830: contig of 1105 bp in length
* 46830 46930: gap of 100 bp
* 46930 48033: contig of 1103 bp in length
* 48033 48133: gap of 100 bp
* 48133 49181: contig of 1048 bp in length
* 49181 49281: gap of 100 bp
* 49281 50381: contig of 1100 bp in length
* 50381 50481: gap of 100 bp
* 50481 51600: contig of 1119 bp in length
* 51600 51700: gap of 100 bp
* 51700 52817: contig of 1117 bp in length
* 52817 52917: gap of 100 bp
* 52917 53967: contig of 1050 bp in length
* 53967 54067: gap of 100 bp
* 54067 55103: contig of 1036 bp in length
* 55103 55203: gap of 100 bp
* 55203 56264: contig of 1061 bp in length
* 56264 56364: gap of 100 bp
* 56364 57469: contig of 1105 bp in length
* 57469 57569: gap of 100 bp
* 57569 58636: contig of 1067 bp in length
* 58636 58736: gap of 100 bp
* 58736 59797: contig of 1061 bp in length
* 59797 59897: gap of 100 bp
* 59897 60979: contig of 1082 bp in length
* 60979 61079: gap of 100 bp
* 61079 62035: contig of 956 bp in length
* 62035 62135: gap of 100 bp
* 62135 63164: contig of 1029 bp in length
* 63164 63264: gap of 100 bp
* 63264 64305: contig of 1041 bp in length
* 64305 64405: gap of 100 bp
* 64405 65427: contig of 1022 bp in length
* 65427 65527: gap of 100 bp
* 65527 66574: contig of 1047 bp in length
* 66574 66674: gap of 100 bp
* 66674 67683: contig of 1009 bp in length
* 67683 67783: gap of 100 bp
* 67783 68876: contig of 1093 bp in length
* 68876 68976: gap of 100 bp
* 68976 69998: contig of 1022 bp in length
* 69998 70098: gap of 100 bp
* 70098 71149: contig of 1051 bp in length
* 71149 71249: gap of 100 bp
* 71249 72339: contig of 1090 bp in length
* 72339 72439: gap of 100 bp

```

```

Query Match      2.0% Score 55; DB 2; Length 81667;
Best Local Similarity 100.0%; Pred. No. 6.7e-19;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1561 TGTGGGACCGCTTAAAGTACTTCAATTCCTAAGAGACTTGATGATTTA 1615
Db 4699 TGTGGGACCGCTTAAAGTACTTCAATTCCTAAGAGACTTGATGATTTA 4753

RESULT 22
AC095790 179667 bp DNA linear HTG 03-OCT-2002
LOCUS Rattus norvegicus clone CH230-9M12, *** SEQUENCING IN PROGRESS ***
DEFINITION 2 unordered pieces.
ACCESSION AC095790
VERSION AC095790.4 GI:22773054
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 179667)
AUTHORS Muzny D,Marie, Metzker M, Lee, Abramson S, Adams C, Alder J,
Allen C, Allen H, Albrooke S, Amin A, Angiano D,
Ayalebech V, Aoyagi A, Ayodeji M, Baca E, Baden H,
Baldwin D, Bandaranaike D, Barber M, Barnstead M, Benahmed F,
Biswal K, Blair J, Blankenburg K, Blyth P, Brown M,
Byrant N, Buhay C, Burch P, Burrell K, Calderon E,
Cardenas V, Carter K, Cavazos I, Caesar H, Center A,
Chacko J, Chavez D, Chen G, Chen R, Chen Y, Chen Z,
Cleveland C, Cockrell R, Cox C, Coyle M, Cree A, D'Souza L,
Davila M, Davis C, Davy-Carroll J, De Ande C, Dederich D,
Delgado O, Denson S, Deramo C, Ding Y, Ding H, Divya K,
Draper H, Dugan-Rocha S, Dunn A, Durbin K, Duval B, Eaves K,
Egan A, Escotto M, Eugene C, Evans C, Falls T, Fan G,
Fernandez S, Finley M, Flagg N, Forbes L, Foster P,
Frazer C, Gabisi A, Ganta R, Garcia A, Garner T, Garza M,
Gebregorjais E, Geer K, Gill R, Grady M, Guerra W, Guevara M,
Gutierrez P, Haaland W, Hamill C, Hamilton K,
Harvey Y, Hawlak P, Hawes A, Henderson N, Hernandez J,
Hernandez R, Hines S, Hladun S, Hume J, Idlebird D, Jackson A,
Hollins B, Howells S, Hulky S, Hume J, Hughes M,
Jackson L, Jacob L, Jiang H, Johnson B, Johnson R, Jolyet A,
Kapathy S, Kelly S, Kelly S, Khan Z, King L, Kovar C,
Kowals C, Kraft C, Lebow H, Levan J, Lewis L, Li Z, Liu J,
Liu J, Liu W, Liu Y, London P, Longacre S, Lopez J,
Lorenshewa L, Louissege H, Losado R, Lu X, Ma J,
Maheshwari M, Mahindaratne M, Mahmoud M, Malloy K, Mangum A,
Mangum B, Mapua P, Martin K, Martin R, Martinez E,
Mawhinney S, McLeod M, McNeill T, Z, Weenen E,
Milosavljevic A, Miner G, Minja E, Montemayor J, Moore S,
Morgan M, Morris K, Morris S, Munidasa M, Murphy M, Nair L,
Nankervis C, Neal D, Newton N, Nguyen N, Norris S,
Nwaekelam O, Okunom G, Olarunsaagun A, Pal S, Parks K,
Pasternak S, Paul H, Perez A, Perez A, Pflanzkoch C,
Plopper F, Poindexter A, Popovic D, Primus B, Pu L,
Puazo M, Quiroz J, Rachlin E, Reeves K, Regier M, Reigh R,

```

```

REFERENCE AUTHORS TITLE JOURNAL
COMMENT
```

Ravelli,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Riles,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scheer,S., Scott,G., Shattam,S., Shen,H., Shetty,U., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajic,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steinle,M., Strong,R., Sutton,A., Svetek,A., Taber,P., Taylor,C., Taylor,T., Thomas,N., Tison,A., Tingey,A., Trejos,Z., Umani,K., Valae,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Zhu,D., von Niederhausen,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstein,G. and Gibbs,R.A.

JOURNAL REFERENCE AUTHORS TITLE JOURNAL COMMENT

Submitted (03-OCT-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Sep 10, 2002 this sequence version replaced gi:21239770.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

--Genome Center--
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

--Project Information--
Center project name: GDUW
Center clone name: CH230-9M12

--Summary Statistics--
Assembly program: Phrap; version 0.990329
Consensus quality: 170329 bases at least Q40
Consensus quality: 171740 bases at least Q30
Consensus quality: 172810 bases at least Q20
Estimated insert size: 192053; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This is a "working draft" sequence. It currently consists of 2 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 89981: contig of 89981 bp in length
* 89982 90081: gap of unknown length
* 90082 179657: contig of 89586 bp in length.
Location/Qualifiers
1. 179657
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-9M12"
90082..91885
/note="wgs_contig"

FEATURES SOURCE misc_feature

BASE COUNT	58000 a	32216 c	32337 g	51139 t	5975 others
ORIGIN					
Query Match	1.8%; Score 50; DB 2; Length 179667;				
Best Local Similarity	100.0%; Pred. No. 4,6e-16;				
Matches	50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Oy	2291	AGCCGAAAGCTGTGGTACAGACTTCATGTGAGCTCATCAACGCCA	2340		
Dn	158228	AGCCGAAAGCTGTGGTACAGACTTCATGTGAGCTCATCAACGCCA	158277		
RESULT_23					
LOCUS	ACI23888				
DEFINITION	Rattus norvegicus clone CH230-43U18, WORKING DRAFT SEQUENCE, 3				
ACCESSION	ACI23888				
VERSION	ACI23888.4				
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.				
SOURCE	Rattus norvegicus (Norway rat)				
ORGANISM	Rattus norvegicus				
REFERENCE	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Ratus.				
AUTHORS	1 (bases 1 to 273260) Mzany,D,Marle, Metzker,M,Lee, Abramzon,S, Adams,C, Alder,J., Allen,C, Allen,H, Alsbrooks,S, Amin,A, Angiano,D., Annalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H., Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F, Bisswal,O, Blair,J, Blankenbury,K, Blyth,P, Brown,M., Bryant,N, Buhay,C, Burch,P, Butrell,K, Calderon,E., Cardenas,J, Carter,K, Cavazos,I, Ceasar,H, Chen,A., Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chtera,Z, Chu,J, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, d'Souza,L, Davis,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D., Delgado,O, Denso,S, Deramo,C, Ding,Y, Dinh,H, Diya,K, Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Evans,K, Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G., Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P., Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garra,M, Gezegeoglis,E, Geier,K, Gill,R, Grady,M, Guerra,M, Guetara,W., Gunaratne,P, Haaland,W, Hambl,C, Hamilton,C, Hamilton,K., Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J., Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hogues,M., Hollins,B, Howells,S, Huliy,S, Hume,J, Idlebird,D, Jackson,A., Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A., Karpachy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovat,C, Kowis,C, Kraft,C,L, Lebow,H, Levay,J, Lewis,L, Li,Z, Liu,J., Liu,T, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J., Lorenhubwa,L, Loulseged,H, Lozada,R.J, Lu,X, Ma,J., Maheshwari,M, Mahindartne,M, Mahmud,M, Malloy,K, Mangum,A., Mangun,B, Mapua,P, Martin,K, Martin,R, Mattheisz,B., Manwhney,S, McLeod,M.P, McNeill,T.Z, Meene,E., Mlosovlevic,A, Miner,G, Minga,E, Montemayor,J, Moore,S., Morgan,M, Morris,K, Morris,S, Mundasa,M, Murphy,M, Nair,L., Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S., Naokeleneh,O, Okuwonu,G, Olarinuyegoon,A, Pal,S, Parks,K., Pasternak,S, Paul,H, Perez,A, Perez,L, Plannkoch,C., Ploppoz,F, Polndexter,A, Popovic,D, Primmus-E, Pu,L-L., Piazom,M, Quitoiz,J, Rachlin,E, Reeves,K, Regier,M.A, Reigh,R, Riley,B, Reilly,M, Ren,Y, Reutter,M, Richards,S, Riggs,F., Rives,C, Rodkey,T, Roja,A, Rose,M, Rose,R, Ruiz,S,J., Sanders,W, Saverly,G, Scherer,S, Scott,G, Shateman,S, Shen,H, Shetty,J, Shvartsbeyn,A, Sisson,I, Sotiller,R, Sosa,J.D., Sneed,A, Sodergren,E, Song,X-Z, Sorelle,R, Sosa,J., Steimle,M, Strong,R, Sutton,A, Svatek,A, Tabor,P, Taylor,C., Taylor,T, Thomas,N, Thomas,S, Tinney,A, Trejos,Z, Usmani,K., Velas,R, Vera,V, Villaseca,D, Waldron,L, Walker,B, Wang,J., Wang,Q, Wang,S, Warren,J, Warren,W, Wei,X, White,F., Williams,G, Willison,R, Wiczzyk,R, Wooden,H., Worley,K., Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X, Zhao,S., Dunn,D, von				

Niederhuser, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G. and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 273260)
 Worley, K.C.
 Direct Submission
 Submitted (03-JUN-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 273260)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (15-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 15, 2002 this sequence version replaced gi:23267611.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with 'N's to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GYRD
 Center clone name: CH230-43J18
 ----- Summary Statistics
 Assembly program: Phrap; version 0.990329
 Consensus quality: 222476 bases at least Q40
 Consensus quality: 224733 bases at least Q30
 Consensus quality: 226257 bases at least Q20
 Estimated insert size: 227079; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
 (see http://www.hgsc.bcm.tmc.edu/doc/Genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of 'N', but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

* 1 268389: contig of 268389 bp in length
 * 268390 268489: gap of unknown length
 * 268490 269678: contig of 1189 bp in length
 * 269679 269778: gap of unknown length
 * 269779 273260: contig of 3482 bp in length.

FEATURES
 source
 1. 273260
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-43J18"
 1. 1061
 /note="wgs_end_extension
 clone_end:17"
 4306 .44451
 /note="clone_boundary
 clone_end:17
 site:

misc_feature
 misc_feature
 misc_feature

misc_feature end_sequence: BH265538"
 58743 .60404
 /note="wgs_contig"
 misc_feature 265291 .266193
 /note="clone_boundary
 clone_end:sp6
 site:
 end_sequence: BH265541"
 BASE COUNT 66654 a 43343 c 44553 g 72796 t 45914 others
 ORIGIN

Query Match 1.8%; Score 50; DB 2; Length 273260;
 Best Local Similarity 100.0%; Pred. No. 4.5e-16;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2291 ACCGAGAGTGTGTACAGACTTCATGCAGCTCATCAACACCCCA 2340
 Db 44201 ACCGAGAGTGTGTACAGACTTCATGCAGCTCATCAACACCCCA 44250

RESULT 24
 MUSC1
 LOCUS
 DEFINITION Mus musculus sema3a mRNA for semaphorin 3A, complete cds.
 ACCESSION D85028
 VERSION D85028.2 GI:18389545
 KEYWORDS collapse-1.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2356)
 Taniguchi, M., Yusa, S., Fujisawa, H., Naruse, I., Saga, S., Mishina, M.
 and Yagi, T.
 Disruption of semaphorin III/D gene causes severe abnormality in
 peripheral nerve projection
 Neuron 19 (3), 519-530 (1997)
 97470885
 MEDLINE 9331345
 PUBMED 2 (bases 1 to 5952)
 Taniguchi, M.
 Direct Submission
 Submitted (02-MAY-1996) Masahiko Taniguchi, Department of
 Biochemistry and Molecular Biology, Graduate School of Medicine,
 The University of Tokyo; Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033,
 Japan (E-mail: taniguchi@m.u-tokyo.ac.jp, Tel: 81-3-5802-2925,
 Fax: 81-3-3813-8732)
 On Jan 28, 2002 this sequence version replaced gi:1313903.

COMMENT
 FEATURES
 source
 1. 5952
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 1. 5952
 /gene="sema3A"
 652. 2970
 /gene="sema3A"
 /codon_start=1
 /product="semaphorin 3A"
 /protein_id="BA19773.1"
 /db_xref="GI:2055316"
 /translation="MGWFTGACLFVGVLLTARANYANGKNNVRLKLSYKEMLESNN
 VITFNGLANSSSYHTFLIDERSRLVYGAOKHISFNLVNIKDPOKIVPVSRYRDE
 CKMAGKDIKESCANFIVLEAYNOTHTVACGTGAFPICTIYEVGHHEVDI FKLDS
 HENGGRGKSPYDPLTLTASLLIDELVSGTADPMGRDPAIFRLGHHNPIRTSDHS
 RWLNDPRTSAHLIPESDNPDDKYVFFRFRNALDGRSGATYARICQITKDPGCH
 RSLVKNKWTFLKARLICSVPNGIDITHFDLQVFLNSKDPKNPVLYGVFTTSSNI
 FKSGAVCMYSMDSVRAVFLGYPARDENYQWVYQGVPRPGTCSKTFGGFST
 KDLPPDVITFARSHAPAMNVPFPIINBPIMIKTIVNYQFTIIVDRVDAEDGQVWF
 IGTGVTVLKVAVSVKPEKTHDLEELVLEMTVPEPTTISMETLSQOOLYIGSDAG
 VAOLPIARCDIYKACACCAIARDPYCAMSSCSRYPRPKXRRTRRDRDIPNGPLTH
 CSDQHDHNGHPSLEERITLYGVNSSTFLFESKSKSQRALVYVQFQRNBRKEEIKR
 GDHIIIRTEGILLRSIQKDSGNYLCAVHEGFMOTLLKTVLEIVTEHLEELHKOD

```

DGDGSKIKEMSSMTPTSQKWRDPMQILNHNNTMDFECCQWKRKORRORPGH
SQGSSNKXKHQDESKGRRNRTHFERAPRSV"
BASE COUNT      1829 a 1203 c 1214 g 1706 t
ORIGIN
Query Match      1.7%; Score 47; DB 10; Length 5952;
Best Local Similarity 100.0%; Pred. No. 2.9e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

118 TTTCAGATTGTTGAACCTCTCTGCGGCACATACGAGAGAG 164
570 TTTCAGATTGTTGAACCTCTCTGCGGCACATACGAGAGAG 616

RESULT 25
AC068065/c
LOCUS DEFINITION Mus musculus clone RP23-30905 strain C57BL6/J, WORKING DRAFT
SEQUENCE, 11 unordered pieces.
AC068065
AC068065.1 GI:7658325
HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 195335)
Bouffard,G.G., Dietrich,N.L., Bagle,M.O., Gupta,J., Ho,S.-L.,
Huang,M.C., Idol,J., Lee-Lin,S.-Q., Maduro,Q.L., Maduro,V.B.,
Maestrian,S.D., McCloskey,J.C., Morse,J., Ojodu,M.A., Pearson,R.,
Stantropop,S., Summers,T.J., Thomas,J.W., Thomas,P.J.,
Tiongson,E.E., Touchman,J.W., Tran,J.T., Vogt,J.L., Walker,M.A.,
Wetherby,K.D. and Green,E.D.
NISC Mouse Sequencing Initiative
Unpublished
2 (bases 1 to 195335)
Green,E.D.
Direct Submision
Submitted (28-APR-2000) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
-----
Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.mouse@nigri.nih.gov
-----
Project Information
Center Project name: YV
Center clone name: 309J05
-----
Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 187134 bases at least Q40
Consensus quality: 188869 bases at least Q30
Consensus quality: 190074 bases at least Q20
Insert size: 190000; agarose-fp
Insert size: 194000; pulse-field-gel
Insert size: 195335; sum-of-contigs
Quality coverage: 8.15x in Q20 bases; agarose-fp
Quality coverage: 7.98x in Q20 bases; pulse-field-gel
Quality coverage: 7.92x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2293: contig of 2293 bp in length
* gap of unknown length

```

```

2294 5070: contig of 2777 bp in length
* gap of unknown length
* 5071 8818: contig of 3748 bp in length
* gap of unknown length
* 8819 16171: contig of 7353 bp in length
* gap of unknown length
* 16172 30056: contig of 13885 bp in length
* gap of unknown length
* 30057 47096: contig of 17040 bp in length
* gap of unknown length
* 47097 68618: contig of 21522 bp in length
* gap of unknown length
* 68619 95138: contig of 26520 bp in length
* gap of unknown length
* 95139 126583: contig of 31445 bp in length
* gap of unknown length
* 126584 156528: contig of 22945 bp in length
* gap of unknown length
* 156529 195335: contig of 38807 bp in length.
Location/Qualifiers
1..195335
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-30905"
/clone_lib="RPMI mouse BAC library 23"
BASE COUNT 61274 a 35333 c 35501 g 62857 t 370 others
ORIGIN
Query Match      1.7%; Score 47; DB 2; Length 195335;
Best Local Similarity 100.0%; Pred. No. 2.4e-14;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

118 TTTCAGATTGTTGAACCTCTCTGCGGCACATACGAGAG 164
27069 TTTCAGATTGTTGAACCTCTCTGCGGCACATACGAGAG 27023

RESULT 26
AC121125/c
LOCUS DEFINITION Mus musculus clone RP23-10F17, *** SEQUENCING IN PROGRESS ***
AC121125
AC121125.4 GI:31581674
HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 265982)
Birtten,B., Nusbaum,C. and Lander,E.
Unpublished
2 (bases 1 to 265982)
Birtten,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Anderson,S., Banna,N., Baetien,V., Bloom,T., Boguslavsky,L.,
Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collins,A.,
Cook,A., Cooke,P., Dearellano,K., Dawar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horton,L., Hulme,W., Riley,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kelle,C., Lakocque,K., Lamezates,R.,
Landers,T., Lehoczy,J., Levine,R., Lindblad-Toh,K., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Margulis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Menus,L.,
Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R.,

```


VERSION L41541.1 GI:18654369
 KEYWORDS collapsin/semaphorinIII.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2319)
 Kimura, T. and Fishman, M.C.
 cDNA sequence of mouse Collapsin/SemaphorinIII
 JOURNAL Unpublished
 FEATURES
 location/Qualifiers
 1..2319
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /dev_stage="adult"
 1..2319
 /note="putative"
 /codon_start=1
 /product="collapsin/semaphorin III"
 /protein_id="AA177611.1"
 /db_xref="GI:18654370"
 /translation="MGWFTGIACLFMGVLLTARANYANGKNVPRILKSYKMLESN
 VITRGLANSSYHTFLDEERSRLVGAADHIFSNLVNIDKQKLVMPVSYRDS
 CKWAGKDLKECANFIKYLEAVNQTHTVACGTGAFHPICTYIEVGHEDNIPFLQDS
 HFENGRKSPYDPLKLTASLLIDGELYSGTADPWGRDPAIFRTLGHHPIRTQDS
 RMLNDPRFISAHLLPESDNPEDDKVYFFERNALIGESKATARIQOLCKDPGGH
 RSLVNMKTTFLKARLICVSPGNGIDTHPELDVFLMNSKDPKNPVYGVFTSSNI
 PKSAVCMYSMDVRVPLGYPYARHDGPNYQVYQGVPPYPRGTCSKTFGGDST
 KDLDDVITFAKSHPMNVPFPIINNPIMIKTDVNYQTOIVDRAVDAGQDVMF
 IGTGVNVLKVVSVPKETWDLLEVMVFEPTTISAMELSTKQOQYIYSTAG
 VAOLPLRCDIYKACAECLARDPYCAMDSSCSRYPYPAKRRTRQDILNGLPLTH
 CSDLQHDNHHGPLEERIITYGVNSTFLKCSPSQALVYQFORNRRSKELIM
 GDHILRTEQGLRLSLQKDSGYLCAVHGHFMOTLLKTLLEVIDTEHLELLHKD
 DGGSKIKEMSSSTPSOKWYTRDPMOLINPNLNTMDEFCQYWKDRKORRPPGH
 SQSSNMKMHQESKKGRNRTHEPERAPRSV"
 BASE COUNT 703 a 513 c 549 g 554 t
 ORIGIN
 Query Match 1.5%; Score 41; DB 10; Length 2319;
 Best Local Similarity 100.0%; Pred. No. 8.1e-11;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2171 CAAACTCTTTTAAGTAACCTGGAAGTCATTGACACAGA 2211
 |||||||
 DB 1975 CAAACTCTTTTAAGTAACCTGGAAGTCATTGACACAGA 2015
 |||||||
 RESULT 29
 MMRNASEMD 2913 bp mRNA linear ROD 08-JUL-1996
 LOCUS Mus musculus mRNA for semaphorin D.
 DEFINITION X85993
 ACCESSION X85993.1 GI:854329
 VERSION X85993.1
 KEYWORDS semaphorin; semd gene.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Puschel, A.W., Adams, R.H. and Betz, H.
 Murine semaphorin D/collapsin is a member of a diverse gene family
 and creates domains inhibitory for axonal extension
 JOURNAL Neuron 14 (5), 941-948 (1995)
 MEDLINE 95267431
 PUBMED 7748561
 REFERENCE 2 (bases 1 to 2913)
 AUTHORS Adams, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (30-MAR-1995) R.H. Adams, Max-Planck-Institute fuer
 Hirnforschung, Deutscherordenstr. 46, D- 60528 Frankfurt, FRG
 COMMENT Sequence overlapping with those under the acc#U02528, L26080,
 L26081 & L26082.

FEATURES
 source
 location/Qualifiers
 1..2913
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="NMRI"
 /db_xref="taxon:10090"
 /clone_1db="Stratagene lambda ZAP"
 /dev_stage="12 days embryo"
 1..2913
 /gene="semd"
 111..2429
 /gene="semd"
 /codon_start=1
 /product="semaphorin D"
 /protein_id="CAA59985.1"
 /db_xref="GI:854330"
 /db_xref="MG1:107558"
 /db_xref="SPTREMBL:Q62180"
 /translation="MGWFTGIACLFMGVLLTARANYANGKNVPRILKSYKMLESN
 VITRGLANSSYHTFLDEERSRLVGAADHIFSNLVNIDKQKLVMPVSYRDS
 CKWAGKDLKECANFIKYLEAVNQTHTVACGTGAFHPICTYIEVGHEDNIPFLQDS
 HFENGRKSPYDPLKLTASLLIDGELYSGTADPWGRDPAIFRTLGHHPIRTQDS
 RMLNDPRFISAHLLPESDNPEDDKVYFFERNALIGESKATARIQOLCKDPGGH
 RSLVNMKTTFLKARLICVSPGNGIDTHPELDVFLMNSKDPKNPVYGVFTSSNI
 PKSAVCMYSMDVRVPLGYPYARHDGPNYQVYQGVPPYPRGTCSKTFGGDST
 KDLDDVITFAKSHPMNVPFPIINNPIMIKTDVNYQTOIVDRAVDAGQDVMF
 IGTGVNVLKVVSVPKETWDLLEVMVFEPTTISAMELSTKQOQYIYSTAG
 VAOLPLRCDIYKACAECLARDPYCAMDSSCSRYPYPAKRRTRQDILNGLPLTH
 CSDLQHDNHHGPLEERIITYGVNSTFLKCSPSQALVYQFORNRRSKELIM
 GDHILRTEQGLRLSLQKDSGYLCAVHGHFMOTLLKTLLEVIDTEHLELLHKD
 DGGSKIKEMSSSTPSOKWYTRDPMOLINPNLNTMDEFCQYWKDRKORRPPGH
 SQSSNMKMHQESKKGRNRTHEPERAPRSV"
 BASE COUNT 888 a 632 c 679 g 714 t
 ORIGIN
 Query Match 1.5%; Score 41; DB 10; Length 2913;
 Best Local Similarity 100.0%; Pred. No. 8e-11;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2171 CAAACTCTTTTAAGTAACCTGGAAGTCATTGACACAGA 2211
 |||||||
 DB 2085 CAAACTCTTTTAAGTAACCTGGAAGTCATTGACACAGA 2125
 |||||||
 RESULT 30
 AC109862_1/c
 WPCOMMENT
 Sequence split into 4 fragments LOCUS AC109862 Accession AC109862
 Fragment Name Begin End
 AC109862_1 1 110000
 AC109862_2 100001 210000
 AC109862_3 200001 310000
 AC109862_4 300001 381240
 Continuation (2 of 4) of AC109862 from base 100001 (AC109862 Rattus norvegicus clone CH2.
 Query Match 1.5%; Score 41; DB 2; Length 110000;
 Best Local Similarity 100.0%; Pred. No. 6.5e-11;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 569 TATATCAGACTCACTTGAAGCTGTGAACGGGGCTTT 609
 |||||||
 DB 99404 TATATCAGACTCACTTGAAGCTGTGAACGGGGCTTT 99364
 |||||||
 RESULT 31
 AC102452 184006 bp DNA linear HTG 09-JUN-2003
 LOCUS Mus musculus clone RP24-272L20, *** SEQUENCING IN PROGRESS ***
 DEFINITION ordered pieces.
 ACCESSION AC102452
 VERSION AC102452.5 GI:31544113
 KEYWORDS HTG; HTGS_PHASE2; HTGS_FULTOP; HTGS_ACTIVEFIN.

SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 184006)
 Biren, B., Nusbaum, C. and Lander, E.
 TITLE Mus musculus, clone RP24-272L20
 JOURNAL Unpublished
 AUTHORS 2 (bases 1 to 184006)
 Biren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barn, N., Baeten, V., Boguslavsky, L., Bouckgalter, B.,
 Brown, A., Camarata, J., Campoliano, A., Chang, J., Chazaro, B.,
 Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A.,
 Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
 Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hages, B., Headford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Kamat, A., Karatas, A., Kells, C., Lachocque, K.,
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
 Maclean, C., MacDonald, P., Major, J., Margolis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrim, J.,
 Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Testfaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Trisillo, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
 JOURNAL Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE 3 (bases 1 to 184006)
 Biren, B., Nusbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, S., Arachch, H.M., Barn, N., Baeten, V., Bloom, T.,
 Boguslavsky, L., Bouckgalter, B., Camarata, J., Chang, J., Choepel, Y.,
 Collamore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
 Diaz, J.S., Dodge, S., Doolley, K., Dorris, L., Erickson, J., Faro, S.,
 Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Graham, L., Grand-Pierre, N., Hages, B., Hagopian, D., Hages, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kells, C., Landers, T., Levine, R.,
 MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Melarim, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
 Rachupka, A., Ramaamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schupback, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
 Talamas, J., Testfaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission
 JOURNAL Submitted (09-JUN-2003) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 COMMENT On Jun 9, 2003 this sequence version replaced gi:11245823.
 All repeats were identified using RepeatMasker:
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: MIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: I18837
 Center clone name: 272_L_20

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 * 1 66973: contig of 66973 bp in length
 * 66974 67073: gap of 100 bp
 * 67074 159784: contig of 92711 bp in length
 * 159785 159884: gap of 100 bp
 * 159885 184006: contig of 24122 bp in length.
 FEATURES
 source
 1. 184006
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="RP24-272L20"
 /clone_11b="RP21-24 Male Mouse BAC"
 /clone_11b="RP21-24 Male Mouse BAC"
 57869 a 34540 c 34334 g 57056 t 207 others

BASE COUNT
 ORIGIN
 57869 a 34540 c 34334 g 57056 t 207 others

Query Match 1.5% Score 41; DB 2; Length 184006;
 Best Local Similarity 100.0%; Pred. No. 6.3e-11;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 2171 CAACTCTTCTTAAGTAACCTGGAAGTCAATGACACAGA 2111
 Db 21139 CAACTCTTCTTAAGTAACCTGGAAGTCAATGACACAGA 21179

RESULT 32
 AC022368/c 207757 bp DNA linear ROD 30-JAN-2001
 LOCUS AC022368
 DEFINITION Mus musculus clone RP23-192D21 strain C57BL6/J, complete sequence.
 ACCESSION AC022368
 VERSION AC022368.2 GI:12597739
 KEYWORDS HTG.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 207757)
 Ayle, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W.,
 Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Grante, S.,
 Guan, X., Gupta, J., Ho, S.-L., Idol, J.R., Karlins, E., Lee-Jin, S.-O.,
 Legaspi, R., Lim, M., Maduro, Q.L., Maduro, V.B., Mastaglio, C.,
 Mastriani, S.D., McCloskey, D.C., McDowell, J., Pearson, R., Prasad, A.,
 Shevchenko, Y., Snyder, B., Stancik, P., Thomas, J.W., Thomas, P.J.,
 Tlonsky, E.E., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A.,
 Wetherby, K.D., Zhang, L.-H. and Green, E.D.
 NISC Comparative Sequencing Initiative

TITLE Unpublished
 JOURNAL 2 (bases 1 to 207757)
 REFERENCE Direct Submission
 JOURNAL Submitted (02-FEB-2000) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 3 (bases 1 to 207757)
 Green, E.D.
 TITLE Direct Submission
 JOURNAL Submitted (30-JAN-2001) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gaithersburg, MD 20877, USA
 COMMENT On Jan 30, 2001 this sequence version replaced gi:16855198.
 Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: http://www.nisc.nih.gov
 Contact: nisc.mouse@nih.gov
 Project Information
 Center project name: Y1
 Center clone name: 192D21

This sequence was finished as follows unless otherwise noted:
all regions were double-stranded, sequenced with an
alternate chemistry, or covered by high quality data
(i.e., phred quality >= 30); an attempt was made to resolve
all sequencing problems, such as compressions and repeats;
all regions were covered by at least one plasmid subclone
or more than one M13 subclone; and the assembly was confirmed
by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of
this clone unless otherwise noted. If there are overlapping
clones, the overlaps are noted in the beginning and end of
the features section.

FEATURES

SOURCE
1..207757
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/clone="RP23-192D21"
/clone_lib="RPCI mouse BAC library 23"
32117..32123
/note="low quality single stranded/single chemistry
region"
misc_feature
32123..32166
/note="single clone coverage"
misc_feature
101083..102298
/note="single clone coverage"
misc_feature
102471..102703
/note="single clone coverage"
misc_feature
102486..102493
/note="low quality single stranded/single chemistry
region"
misc_feature
151895..152199
/note="single clone coverage"
misc_feature
162474..162566
/note="single clone coverage"
misc_feature
190297..190474
/note="single clone coverage"
BASE COUNT 65923 a 38017 c 37853 g 65964 t
ORIGIN

Query Match 1.5%; Score 41; DB 10; Length 207757;
Best Local Similarity 100.0%; Pred. No. 6.2e-11;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2171 CAAACTCTTCTTAAGTAACCTGGAAGTCATTGACACAGA 2211
Db 77684 CAAACTCTTCTTAAGTAACCTGGAAGTCATTGACACAGA 77644

RESULT 33
AC068065 195335 bp DNA linear HTG 28-APR-2000
LOCUS Mus musculus clone RP23-309J5 strain C57BL6/J, WORKING DRAFT
DEFINITION
SEQUENCE 11 unordered pieces.
AC068065
VERSION AC068065.1 GI:7658325
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 195335)

Beckstrom-Stenberg, S.M., Benjamin, B., Blakeley, R.W.,
Bouffard, G.G., Dietrich, N.L., Eagle, W.O., Gupta, J., Ho, S.-L.,
Huang, M.C., Idol, J., Lee-Ian, S.-Q., Maduro, O.L., Maduro, V.B.,
Mastrian, S.D., McCloskey, J.C., Morse, E., Ojodu, M.A., Pearson, R.,
Stancirpop, S., Summers, T.J., Thomas, J.W., Thomas, P.J.,
Tiongson, E.E., Touchman, J.W., Tran, J.T., Vogt, J.L., Walker, M.A.,
Wetherby, K.D. and Green, E.D.
TITLE NISC Mouse Sequencing Initiative
JOURNAL Unpublished

REFERENCE 2 (bases 1 to 195335)
AUTHORS Green, E.D.
TITLE Direct SubMISSION
JOURNAL Submitted (28-APR-2000) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA

COMMENT

Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.mouse@nhgri.nih.gov
Project Information
Center project name: YV
Center clone name: 309J05

Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 187134 bases at least Q40
Consensus quality: 188869 bases at least Q30
Consensus quality: 190074 bases at least Q20
Insert size: 190000; agarose-fp
Insert size: 194000; pulse-field-gel
Insert size: 195335; sum-of-contigs
Quality coverage: 8.15x in Q20 bases; agarose-fp
Quality coverage: 7.98x in Q20 bases; pulse-field-gel
Quality coverage: 7.92x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 11 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2293: contig of 2293 bp in length
* gap of unknown length
* 2294 5070: contig of 2777 bp in length
* gap of unknown length
* 5071 8818: contig of 3748 bp in length
* gap of unknown length
* 8819 16171: contig of 7353 bp in length
* gap of unknown length
* 16172 30056: contig of 13885 bp in length
* gap of unknown length
* 30057 47096: contig of 17040 bp in length
* gap of unknown length
* 47097 68618: contig of 21522 bp in length
* gap of unknown length
* 68619 95138: contig of 26520 bp in length
* gap of unknown length
* 95139 126583: contig of 31445 bp in length
* gap of unknown length
* 126584 156528: contig of 29945 bp in length
* gap of unknown length
* 156529 195335: contig of 38807 bp in length.

FEATURES

SOURCE
1..195335
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL6/J"
/db_xref="taxon:10090"
/clone="RP23-309J5"
/clone_lib="RPCI mouse BAC library 23"
BASE COUNT 61274 a 35333 c 35501 g 62857 t 370 others
ORIGIN

Query Match 1.2%; Score 33; DB 2; Length 195335;
Best Local Similarity 100.0%; Pred. No. 2.3e-06;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 310 AGAATGTTGGAATCAACATGATGATCACTT 342
|||||

DB 176419 AGAATGTTGATCCACATGATGATCTTT 176451

RESULT 34
AC121585 195311 bp DNA linear ROD 14-DEC-2002
LOCUS Mus musculus chromosome 5 clone RP23-265M4, complete sequence.
DEFINITION AC121585
AC121585 GI:26801335
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 195311)
McPherson, J.D. and Waterston, R.H.
The sequence of Mus musculus clone
Unpublished
JOURNAL 2 (bases 1 to 195311)
REFERENCE 2 (bases 1 to 195311)
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submision
JOURNAL Submitted (20-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 195311)
REFERENCE 3 (bases 1 to 195311)
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submision
JOURNAL Submitted (04-SEP-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
4 (bases 1 to 195311)
REFERENCE 4 (bases 1 to 195311)
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submision
JOURNAL Submitted (14-DEC-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT On Dec 14, 2002 this sequence version replaced gi:22711832.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: M_BA0265M04

Location/Qualifiers
1. 195311
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="5"
/clone="RP23-265M4"
BASE COUNT 63171 a 38828 c 37957 g 55355 t
ORIGIN

Query Match 1.2%; Score 32; DB 10; Length 195311;
Best Local Similarity 100.0%; Pred. No. 8.6e-06;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 311 GAAATGTTGATCCACATGATGATCTTT 342
|||||
Db 85566 GAAATGTTGATCCACATGATGATCTTT 85587

RESULT 35
AX359936 4661 bp DNA linear PAT 13-FEB-2002
LOCUS AX359936
DEFINITION Sequence 10 from Patent WO0202783.
ACCESSION AX359936
VERSION AX359936.1 GI:18675575
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1
TITLE
JOURNAL
AUTHORS
TITLE
JOURNAL

AUTHORS Bleck, G.T.
TITLE Expression vectors
JOURNAL Patent: WO 0202783-A 10 10-JAN-2002;
Gala Design, Inc. (US)
FEATURES
source
1. 4661
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="synthetic"
BASE COUNT 1257 a 1096 c 1161 g 1147 t
ORIGIN

Query Match 1.1%; Score 31; DB 6; Length 4661;
Best Local Similarity 100.0%; Pred. No. 4e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATGATGTTAACACGCTT 31
|||||
Db 4041 AATCTTTATTTATGATGTTAACACGCTT 4011

RESULT 36
AX382150 4661 bp DNA linear PAT 18-MAR-2002
LOCUS AX382150
DEFINITION Sequence 10 from Patent WO0202738.
ACCESSION AX382150
VERSION AX382150.1 GI:19576959
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1
AUTHORS Brenel, R.D., Miller, L.U., Bleck, G.T. and York, D.
TITLE Host cells containing multiple integrating vectors
JOURNAL Patent: WO 0202738-A 10 10-JAN-2002;
Gala Design, Inc. (US)
FEATURES
source
1. 4661
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
BASE COUNT 1257 a 1096 c 1161 g 1147 t
ORIGIN

Query Match 1.1%; Score 31; DB 6; Length 4661;
Best Local Similarity 100.0%; Pred. No. 4e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATGATGTTAACACGCTT 31
|||||
Db 4041 AATCTTTATTTATGATGTTAACACGCTT 4011

RESULT 37
SYNMLP1N3 6620 bp DNA linear SYN 05-JAN-2001
LOCUS SYNMLP1N3
DEFINITION Retroviral vector pLNCX, complete sequence.
ACCESSION M28247
VERSION M28247.1 GI:208846
KEYWORDS genome; neomycin phosphotransferase; retroviral vector.
SOURCE Retroviral vector pLNCX
ORGANISM Retroviral vector pLNCX
REFERENCE 1 (bases 1 to 6620)
AUTHORS Miller, A.D.
TITLE Direct Submision
JOURNAL Submitted (22-SEP-1989) A.D. Miller, Program in Molecular Medicine,
Fred Hutchinson Cancer Research Center, Seattle, WA 98104, USA
2 (bases 1 to 6620)
REFERENCE 2 (bases 1 to 6620)
AUTHORS Miller, A.D. and Rosman, G.J.
TITLE Improved retroviral vectors for gene transfer and expression
JOURNAL Biotechniques 7 (9), 980-982 (1989)

FEATURES
MEDLINE 90211945
PUBMED 2631796
SOURCE

Location/Qualifiers
1..6620
/organism="Retroviral vector pLNCX"
/mol_type="genomic DNA"
/db_xref="taxon:147133"
/note="derived from Moloney murine leukemia virus"
1447145
/note="mouse DNA end-Mo-MuSV DNA start"
145..733
/note="Mo-MuSV 5' long terminal repeat"
803..1612
/note="extended packaging signal"
113371134
/note="Mo-MuSV DNA end-Mo-MuLV DNA start"
1193..1195
/note="gag ATG start codon to TAG stop codon"
161671617
/note="Mo-MuLV DNA end-Tn5 DNA start"
1656..2450
/note="neomycin phosphotransferase"
/codon_start=1
/transl_table=1
/protein_id="AA72064.1"
/db_xref="GI:208847"
/translation="MIEQDGLHAGSPAAWVERLFGYDMAOQTICGDAAVPRLSAQR
PVEFKTDISALNELDEBARLSMTATGVCACAVDVTTEAGRDMLLGEVGGDL
LHSLAPAEKVSIMADAKRLHTLDPATCPDPAKAKHIERAKTRMEGLVDQDLDB
EHQGLAPAELEFARLKARMPDEDLVTHGDCLEINWENGRFSGFTIDCGRLGVADRY
QDILATRDIDIEELGEMADRFVLVYGIAPDSORIAFYRLDEFF"
279972800
/note="Tn5 DNA end-human cytomegalovirus DNA start"
2800..3617
/note="human cytomegalovirus immediate early promoter"
36173618
/note="human cytomegalovirus DNA end-Mo-MuLV DNA start"
3666..4259
/note="Mo-MuLV 3' long terminal repeat"
432874329
/note="Mo-MuLV DNA end-plasmid pBR322 DNA start"
1807 c 1692 g 1593 t

BASE COUNT 1528 a 1807 c 1692 g 1593 t
ORIGIN

Query Match 1.1%; Score 31; DB 12; Length 6620;
Best Local Similarity 100.0%; Pred. No. 3.9e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 AATCTTTATTTATCGATGTTAACAGCTT 31
|||||
Db 3639 AATCTTTATTTATCGATGTTAACAGCTT 3609

RESULT 38
SYNMOV2/c 6837 bp DNA linear SYN 27-APR-1993
LOCUS Moloney murine leukemia virus retroviral vector pLNCX, complete
DEFINITION
ACCESSION M64754
VERSION M64754.1 GI:208857
KEYWORDS cloning vector; histidinol dehydrogenase; retroviral vector.
SOURCE unidentified cloning vector
ORGANISM unidentified cloning vector.
REFERENCE 1 (bases 1 to 6837)
AUTHORS Stockschlaeder, M.A., Storb, R., Osborne, W.R. and Miller, A.D.
TITLE L-histidinol provides effective selection of
retrovirus-vector-transduced keratinocytes without impairing their
proliferative potential
Hum. Gene Ther. 2 (1), 33-39 (1991)
JOURNAL Hum. Gene Ther. 2 (1), 33-39 (1991)
MEDLINE 91322161
PUBMED 1650586
COMMENT Original source, text: Synthetic, DNA.

FEATURES
SOURCE

Location/Qualifiers
1..6837
/organism="unidentified cloning vector"
/mol_type="genomic DNA"
/db_xref="taxon:45196"
175..763
/note="Mo-MuSV 5' long terminal repeat"
833..1642
/note="extended packaging signal"
1223..1225
/note="gag ATG start codon to TAG stop codon"
1672..2976
/EC_number="1.1.1.23"
/codon_start=1
/transl_table=1
/product="histidinol dehydrogenase"
/protein_id="AA73024.1"
/db_xref="GI:208858"
/translation="MSFNTLIDMNSCSPEOQALITRPASISDSITRTVSDILNFK
TRGDDALREYSAPKDKTEVTALRTPEELIAAGARLSDELKOMTAIVAKNIEFHSQ
TLAPVDVETQPGVRCQOVTSPVSGVITPGSAPLSTETMLATTPARIAGCQKVVC
SPPIADBITIYAAQLCCVQGEIENVGQAQATIALAFSGESVPKDKIFGPNATFTEK
RQVSQRUDGAIDMPAGSEVLVYADSGATPDVVASDLSQAHEGPSQVTLTPDAD
IARKVAAVAVROLAELPRADTARQALSASRLIVTKDIAQCVAISNOYGPRLIQTEN
ARDLVDAITSGAGSVFLGDMSPESAGDVASGTNNHLYPYGYATCSISGLADFOKRMV
OELSKAGFSALASTIETLAAERLTAHNAVTLRVNALKEQA"
3017..3834
/note="human cytomegalovirus early promoter"
3883..4476
/note="Mo-MuLV 3' long terminal repeat"
1877 c 1722 g 1628 t

BASE COUNT 1610 a 1877 c 1722 g 1628 t
ORIGIN

Query Match 1.1%; Score 31; DB 12; Length 6837;
Best Local Similarity 100.0%; Pred. No. 3.9e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 AATCTTTATTTATCGATGTTAACAGCTT 31
|||||
Db 3856 AATCTTTATTTATCGATGTTAACAGCTT 3826

RESULT 39
AR302096/c 7160 bp DNA linear PAT 12-JUN-2003
LOCUS Sequence 5 from patent US 6541197.
DEFINITION
ACCESSION AR302096
VERSION AR302096.1 GI:31690130
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7160)
AUTHORS Link, C.J. Jr., Levy, J.P., Wang, S. and Seregina, T.
TITLE Vehicles for stable transfer of green fluorescent protein gene and
methods of use for same
Patent: US 6541197-A 5 01-APR-2003;
JOURNAL Location/Qualifiers
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

Query Match 1.1%; Score 31; DB 6; Length 7160;
Best Local Similarity 100.0%; Pred. No. 3.9e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 AATCTTTATTTATCGATGTTAACAGCTT 31
|||||
Db 4179 AATCTTTATTTATCGATGTTAACAGCTT 4149

RESULT 40

AR302097/c AR302097 7235 bp DNA linear PAT 12-JUN-2003
 LOCUS Sequence 6 from patent US 6541197.
 DEFINITION AR302097
 ACCESSION AR302097
 VERSION AR302097.1 GI:31690131
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE
 1 (bases 1 to 7235)
 Link, C.J. Jr., Levy, J.P., Wang, S. and Seregina, T.
 Authors Vehicles for stable transfer of green fluorescent protein gene and
 Title methods of use for same
 Journal Patent: US 6541197-A 6 01-APR-2003;
 Features Location/Qualifiers
 source 1..7235
 /organism="unknown"
 BASE COUNT 1662 a 2088 c 1880 g 1605 t
 ORIGIN

Query Match 1.1%; Score 31; DB 6; Length 7235;
 Best Local Similarity 100.0%; Pred. No. 3.9e-05;
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATGATGTTAACAAGCTT 31
 Db 4254 AATCTTTATTTATGATGTTAACAAGCTT 4224

RESULT 41
 AX591134 29 bp DNA linear PAT 27-JAN-2003
 LOCUS AX591134
 DEFINITION Sequence 7 from Patent WO02086157.
 ACCESSION AX591134
 VERSION AX591134.1 GI:27949649
 KEYWORDS
 SOURCE unidentified
 ORGANISM unidentified
 REFERENCE 1
 Latil, A., Cusano, O. and Algate-Genin, M.
 Authors Use of semaphorin 3a for diagnosing and treating cancer, especially
 Title prostate cancer
 Journal Patent: WO 02086157-A 7 31-OCT-2002;
 Features Location/Qualifiers
 source 1..29
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"
 /note="Amorce utilisee pour supprimer le codon stop de la
 traduction de l'ADNc de la semaphorine-3a (amorce 5')"

BASE COUNT 9 a 6 c 5 g 9 t
 ORIGIN

Query Match 1.1%; Score 29; DB 6; Length 29;
 Best Local Similarity 100.0%; Pred. No. 0.00073;
 Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2155 GGAACATGGGTCATCAACAACCTTTT 2183
 Db 1 GGAACATGGGTCATCAACAACCTTTT 29

RESULT 42
 AF086761 3148 bp mRNA linear VRT 16-JUL-1999
 LOCUS AF086761
 DEFINITION Dario rerio semaphorin 3a mRNA, complete cds.
 ACCESSION AF086761
 VERSION AF086761.1 GI:5499716
 KEYWORDS
 SOURCE Dario rerio (zebrafish)
 ORGANISM Dario rerio

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Osteichthyes;
 Cypriniformes; Cyprinidae; Danio.
 REFERENCE
 1 (bases 1 to 3148)
 Yee, C.S., Chandrasekhar, A., Halloran, M.C., Shoji, W., Warren, J.T.
 and Kuwada, J.Y.
 Authors Molecular cloning, expression, and activity of zebrafish semaphorin
 Title 31a
 Journal Brain Res. Bull. 48 (6), 581-593 (1999)
 MEDLINE 99313409
 PUBMED 10386838
 REFERENCE
 2 (bases 1 to 3148)
 Yee, C.S., Chandrasekhar, A., Halloran, M.C., Shoji, W., Warren, J.T.
 and Kuwada, J.Y.
 Authors Direct Submision
 Title Submitted (24-AUG-1998) Biology, University of Michigan, 830 N.
 Journal University, Ann Arbor, MI 48109-1048, USA
 Features Location/Qualifiers
 source 1..3148
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 358..2940
 /codon_start=1
 /product="semaphorin 31a"
 /protein_id="AAD43964.1"
 /db_xref="GI:5499717"

CDS
 358..2940
 /codon_start=1
 /product="semaphorin 31a"
 /protein_id="AAD43964.1"
 /db_xref="GI:5499717"

BASE COUNT 864 a 804 c 800 g 680 t
 ORIGIN

Query Match 1.0%; Score 27; DB 5; Length 3148;
 Best Local Similarity 100.0%; Pred. No. 0.0078;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 500 GATGATGCAAGTGGGCTGGAAGAAGAC 526
 Db 661 GATGATGCAAGTGGGCTGGAAGAAGAC 687

RESULT 43
 AX107928 9729 bp DNA linear PAT 30-APR-2001
 LOCUS AX107928
 DEFINITION Sequence 56 from Patent WO0125466.
 ACCESSION AX107928
 VERSION AX107928.1 GI:13923311
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 REFERENCE 1
 Slingsby, J., Kingsman, S.M., Rohll, J.O. and Slade, A.O.
 Authors Producer cell for the production of retroviral vectors
 Title Patent: WO 0125466-A 56 12-APR-2001;
 Journal Oxford Biomedica (UK) Limited (GB)
 Features Location/Qualifiers
 source 1..9729
 /organism="synthetic construct"

/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic sequence"
BASE COUNT 2270 a 2657 c 2540 g 2262 t
ORIGIN

Query Match 1.0%; Score 27; DB 6; Length 9729;
Best Local Similarity 100.0%; Pred. No. 0.0073;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGTTAACAA 27
|||||
Db 5278 AATCTTTATTTATCGATGTTAACAA 5252

RESULT 44
AX107927/c 12473 bp DNA linear PAT 30-APR-2001
LOCUS
DEFINITION Sequence 55 from Patent WO0125466.
ACCESSION AX107927
VERSION AX107927.1 GI:13923310
KEYWORDS
SOURCE synthetic construct
ORGANISM artificial sequences.

REFERENCE 1
AUTHORS Slingsby, J., Kingsman, S.M., Rohlf, J.O. and Slade, A.O.
TITLE Producer cell for the production of retroviral vectors
JOURNAL Patent: WO 0125466-A 55 12-APR-2001;
Oxford Biomedica (UK) Limited (GB)
FEATURES
SOURCE 1. 12473
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="Synthetic sequence"

BASE COUNT 2841 a 3434 c 3357 g 2841 t
ORIGIN

Query Match 1.0%; Score 27; DB 6; Length 12473;
Best Local Similarity 100.0%; Pred. No. 0.0072;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGTTAACAA 27
|||||
Db 5278 AATCTTTATTTATCGATGTTAACAA 5252

RESULT 45
AX591129/c 25 bp DNA linear PAT 27-JAN-2003
LOCUS
DEFINITION Sequence 2 from Patent WO02086157.
ACCESSION AX591129
VERSION AX591129.1 GI:27949644
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Latil, A., Cussenot, O. and Alzarde-Genin, M.
TITLE Use of semaphorin 3a for diagnosing and treating cancer, especially prostate cancer
JOURNAL Patent: WO 02086157-A 2 31-OCT-2002;
Urogene Societe anonyme (FR)
FEATURES
SOURCE 1. 25
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
/note="Amorce d'amplification du gene de la semaphorine-3a(lower primer) choisie avec l'assistance des programmes informatiques Oligo4 (National Biosciences, Plymouth, MN) et Primer express (Perkin-Elmer Applied

Bioesystems, Foster city, CA)"
BASE COUNT 7 a 7 c 4 g 7 t
ORIGIN

Query Match 0.9%; Score 25; DB 6; Length 25;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1556 ACAGTGTGGACCGTTCTTAAG 1580
|||||
Db 25 ACAGTGTGGACCGTTCTTAAG 1

RESULT 46
AX207158 1400 bp DNA linear PAT 30-AUG-2001
LOCUS
DEFINITION Sequence 5 from Patent WO0155455.
ACCESSION AX207158
VERSION AX207158.1 GI:15394946
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1
AUTHORS Jin, S.
TITLE Resistance sequences and uses thereof
JOURNAL Patent: WO 0155455-A 5 02-AUG-2001;
Millennium Pharmaceuticals, Inc. (US) ; Jin, Shengfang (US)
FEATURES
SOURCE 1. 1400
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

BASE COUNT 437 a 313 c 278 g 372 t
ORIGIN

Query Match 0.9%; Score 25; DB 6; Length 1400;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 TAACAGCTTAGTATGATGCCAC 46
|||||
Db 29 TAACAGCTTAGTATGATGCCAC 53

RESULT 47
AC098444/c 244749 bp DNA linear HTG 13-NOV-2002
LOCUS
DEFINITION Rattus norvegicus clone CH230-2G15, *** SEQUENCING IN PROGRESS ***
ACCESSION AC098444
VERSION AC098444.4 GI:24941288
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

REFERENCE 1
AUTHORS Muzny, D.M., Adams, C., Adio-Oduola, B., Ali-osman, F.R., Allen, C., Alibrooke, S.L., Amaralunga, H.C., Arc, J.R., Ayala, M., Banks, T., Barbara, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowie, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Burch, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Caron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhury, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M., Falis, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,

Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
 Gorrell, J.H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K.,
 Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
 Hernandez, O., Hodgson, A., Hognes, M., Holloway, C., Hollins, B.,
 Homai, F., Howard, S., Huber, J., Hulyk, S., Hume, J., Jackson, L.E.,
 Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
 Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C.,
 Kravtsov, J., Kureishi, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
 Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W., Louieged, H.,
 Lozdo, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
 Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, R.,
 Massey, E., Mawhinney, E., McLeod, M.P., Meador, M., Mei, G., Metzger, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, N., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokenko, S., Ogih, M., Okunolu, G.,
 Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rivers, M., Rojas, A., Rojudo, K., Rolfe, M., Ruiz, S., Savery, G.,
 Scherer, S., Scott, G., Shen, H., Shoshari, N., Sisson, I.,
 Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
 Usmani, K., Vaequez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (13-NOV-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 13, 2002 this sequence version replaced gi:23664578.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rac/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by fixed gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: TUBB

Center clone name: CH230-2G15

----- Summary Statistics

Assembly program: Phrap, version 0.990329

Consensus quality: 231469 bases at least Q40

Consensus quality: 234436 bases at least Q30

Consensus quality: 236097 bases at least Q20

Estimated insert size: 242749; sum-of-contigs estimation

Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

FEATURES

source

1. 244749

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-2G15"

1. 1633

/note="wgs_end_extension"

clone_end:Sp6"

1885..3960

/note="wgs_end_extension"

clone_end:Sp6"

complement(5662..6488)

/note="clone_boundary"

clone_end:Sp6"

site:

end sequence: BH287303"

complement(239764..240641)

/note="clone_boundary"

clone_end:T7"

site:

end sequence: BH287302"

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches

25; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

QY

1791

GAGACCTTACGTGCTGGGATG 1815

Db

228396

GAGACCTTACGTGCTGGGATG 228372

RESULT 48

AC096525

LOCUS

DEFINITION

AC096525

AC096525

VERSION

AC096525.7

AC096525

KEYWORDS

HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

SOURCE

Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 250672)

REFERENCE

AUTHORS

Muzny, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J.,

Allen, C., Allen, H., Alibrooks, S., Amin, A., Argüello, D.,

Ayalele, V., Ayagi, A., Ayodeji, M., Baca, E., Baden, H.,

Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,

Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,

Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,

Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A.,

Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,

Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,

Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

Drapper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,

Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G.,

Fernandez, S., Finley, M., Flaeg, N., Forbes, L., Foster, M., Foster, P.,

Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,

Gebregorgis, B., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idledbid, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Kapachy, S., Kelly, S., Kelly, S., Khan, Z., Khan, L., Kovar, C., Kowals, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lourenshaw, L., Louised, H., Lozano, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinley, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundaya, M., Murphy, M., Nair, L., Narkervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokenkeme, O., Okumu, G., Olariunbasoon, A., Pal, S., Parks, K., Patel, S., Paul, H., Perez, A., Perez, L., Pflannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, B., Pu, L., L., Puzo, M., Quiroz, J., Rachlin, B., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajic, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, D., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

TITLE

Unpublished

REFERENCE

2 (bases 1 to 250672)

AUTHORS

Worley, K. C.

TITLE

Submitted (18-SEP-2001)

JOURNAL

Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE

3 (bases 1 to 250672)

AUTHORS

Rat Genome Sequencing Consortium.

TITLE

Submitted (10-MAY-2003)

JOURNAL

Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE

On May 10, 2003 this sequence version replaced gi:24959330.

AUTHORS

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas

TITLE

(http://www.hgsc.bcm.tmc.edu/projects/rat/) Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

JOURNAL

Genome Center

REFERENCE

Center: Baylor College of Medicine

AUTHORS

Center code: BCM

TITLE

Web site: http://www.hgsc.bcm.tmc.edu/

JOURNAL

Contact: hgsc-help@bcm.tmc.edu

REFERENCE

Project Information

AUTHORS

Center project name: CGBP

TITLE

Center clone name: CH230-84N24

JOURNAL

Summary Statistics

REFERENCE

Assembly program: Atlas 3.0;

AUTHORS

Consensus quality: 229622 bases at least Q40

TITLE

Consensus quality: 232300 bases at least Q30

Consensus quality: 234233 bases at least Q20
Estimated insert size: 241730; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 237667: contig of 237667 bp in length
* 237668 237667: gap of unknown length
* 237768 247530: contig of 9763 bp in length
* 247531 247630: gap of unknown length
* 247631 250672: contig of 3042 bp in length.

FEATURES

source

1. 250672
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-84N24"

misc_feature

1. 2033
/note="wgs contig"

misc_feature

237768..238784
/note="wgs contig"

BASE COUNT 72811 a 44292 c 43115 g 74893 t 15561 others
ORIGIN

Query Match 0.9%; Score 25; DB 2; Length 250672;
Best Local Similarity 100.0%; Pred. No. 0.084;
Matches 25; Conservative 0; Mismatches 0; Gaps 0;

Qy

1791 GAGACCTTACTGCTGGGATG 1815

Db

170733 GAGACCTTACTGCTGGGATG 170757

RESULT 49

AX591128 24 bp DNA linear PAT 27-JAN-2003

LOCUS

AX591128 Sequence 1 from Patent WO02086157.

DEFINITION

AX591128

ACCESSION

AX591128.1 GI:27949643

VERSION

AX591128.1 GI:27949643

KEYWORDS

unidentified
unclassified

SOURCE

unclassified

ORGANISM

unclassified

REFERENCE

1. Latil, A., Cussenot, O. and Algate-Genin, M.
Use of semaphorin 3a for diagnosing and treating cancer, especially prostate cancer
Patent: WO 02086157-A 1 31-OCT-2002;

JOURNAL

Urogene Societe anonyme (PR)

FEATURES

Location/Qualifiers

source

1. 24
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"
/note="Name of amplification du gene de la semaphorine-3A (upper primer) choisie avec l'assistance des programmes informatiques Oligo4 (National BioSciences, Plymouth, MN) et Primer express (Perkin-Elmer Applied Biosystems, Foster city, CA)"

BASE COUNT

8 a 7 c 4 g 5 t

ORIGIN

Query Match

0.9%; Score 24; DB 6; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.53;
Matches 24; Conservative 0; Mismatches 0; Gaps 0;

JOURNAL Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:23664931.

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GHDG
Center clone name: CH230-155C17

----- Summary Statistics

Assembly program: Atlas 3.0;
Consensus quality: 215727 bases at least Q40
Consensus quality: 218302 bases at least Q30
Consensus quality: 219785 bases at least Q20
Estimated insert size: 233913; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 3 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 218040: contig of 218040 bp in length
* 218041 218140: gap of unknown length
* 218141 222614: contig of 4474 bp in length
* 222615 222714: gap of unknown length
* 222715 225005: contig of 2291 bp in length.
Location/Qualifiers

FEATURES
SOURCE
1. 225005

/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-155C17"

misc_feature

/note="wgs contig"
complement(215871..216597)
/note="clone_boundary"
clone_end:T7

misc_feature

end_sequence:BH364008"
218141..219872
/note="wgs_end_extension"
clone_end:T7"

misc_feature

219923..222614
/note="wgs_end_extension"
clone_end:T7"

BASE COUNT 66589 a 44892 c 44710 g 64099 t 4715 others

ORIGIN

Query Match 0.9%; Score 24; DB 2; Length 225005;
Best Local Similarity 100.0%; Pred. No. 0.31;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1075 TTGCTCAGTCCAGTCCCAATGG 1098
Db 127822 TTGCTCAGTCCAGTCCCAATGG 127845

RESULT 52

AX591132 29 bp DNA linear PAT 27-JAN-2003
LOCUS Sequence 5 from Patent WO02086157.

DEFINITION AX591132

ACCESSION AX591132.1 GI:27949647

VERSION AX591132.1

KEYWORDS

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

SOURCE

1. 29

/organism="unidentified"

/mol_type="genomic DNA"

/db_xref="taxon:32644"

/note="Amorce d'amplification du gene de la semaphorine-3A(amorce 5')"

semaphorine-3A(amorce 5')"

Location/Qualifiers

1. 29

/organism="unidentified"

/mol_type="genomic DNA"

/db_xref="taxon:32644"

/note="Amorce d'amplification du gene de la semaphorine-3A(amorce 5')"

semaphorine-3A(amorce 5')"

Location/Qualifiers

1. 29

/organism="unidentified"

/mol_type="genomic DNA"

/db_xref="taxon:32644"

/note="Amorce d'amplification du gene de la semaphorine-3A(amorce 5')"

semaphorine-3A(amorce 5')"

Location/Qualifiers

1. 29

/organism="unidentified"

RESULT 53

AX591132 29 bp DNA linear PAT 27-JAN-2003
LOCUS Sequence 5 from Patent WO02086157.

DEFINITION AX591132

ACCESSION AX591132.1 GI:27949647

VERSION AX591132.1

KEYWORDS

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

SOURCE

1. 29

/organism="unidentified"

/mol_type="genomic DNA"

/db_xref="taxon:32644"

/note="Amorce d'amplification du gene de la semaphorine-3A(amorce 5')"

semaphorine-3A(amorce 5')"

Location/Qualifiers

1. 29

/organism="unidentified"

/mol_type="genomic DNA"

/db_xref="taxon:32644"

/note="Amorce d'amplification du gene de la semaphorine-3A(amorce 5')"

semaphorine-3A(amorce 5')"

Location/Qualifiers

1. 29

/organism="unidentified"

/mol_type="genomic DNA"

/db_xref="taxon:32644"

/note="Amorce d'amplification du gene de la semaphorine-3A(amorce 5')"

semaphorine-3A(amorce 5')"

Location/Qualifiers

1. 29

/organism="unidentified"

gene
CDS
/dev_stage="adult"
1.1822
/gene="MMSTK1"
184..1221
/note="putative serine/threonine kinase"
/codon_start=1
/product="STK-1"
/db_xref="GI:1322144"
/translation="MAKONAYMPYSGKTSQSGINTLSORVIREPATSLALAVN
SNSQTAAPGQKLAENKSGSTSGQSNKQPTTINFEIGPVGKRGNYTLAREK
KSRFVALKILEKSGQIEKGEVHQLELEIQAHKGNPILQYNYFYQORLYILE
YAPRGELAYELQKSRTFDEORTATIMEELSDALYCHKKKVIRHDIKPNLLGQGE
LKADGMSVHAAPSIRKTMCGTLDYLPPEMIEGRHNMVMDLMCGVLCYELMGNP
PPSPSHSEYRIYAVDLPSPSVPSGADDISKLKKNPMQRLPLAVAAHPWRA
NSRRVLPSPAL"
1794..1799
/gene="MMSTK1"
1822
/gene="MMSTK1"
polyA_signal
polyA_site
BASE COUNT 435 a 497 c 449 g 441 t
ORIGIN
Query Match 0.8%; Score 23; DB 10; Length 1822;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 57 GTGACCCACGCGTCCGGAGTA 79
|||||
Db 22 GTGACCCACGCGTCCGGAGTA 44
RESULT 54
AY030051 2325 bp mRNA linear VRT 01-MAY-2002
LOCUS Xenopus laevis semaphorin 3A mRNA, complete cds.
DEFINITION
ACCESSION AY030051
VERSION AY030051.1 GI:20384676
KEYWORDS
SOURCE
ORGANISM
Xenopus laevis (African clawed frog)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae;
Xenopodinae; Xenopus.
1 (bases 1 to 2325)
Xenopus laevis
Tannahill, D., Nielsen, J. and Regan, A. G.
Xenopus Semaphorin 3A
Unpublished
2 (bases 1 to 2325)
Tannahill, D., Nielsen, J. and Regan, A. G.
Direct Submission
Submitted (11-APR-2001) Anatomy, University of Cambridge, Downing
Street, Cambridge CB2 3DY, UK
location/Qualifiers
1..2325
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/feature_type="head"
1..2325
/note="class 3 secreted semaphorin; axonal guidance
protein"
/product="semaphorin 3A"
/codon_start=1
/protein_id="AAK38166.1"
/db_xref="GI:20384677"
/translation="MGSILTWIAFLFGLVSLITRLNCONVKNVPRILRLSYKEMVESNN
LITFNGWNSSTYNTFLDEERGLVYGAKHI FSNLANVKEPQKILAMPYINLRDE
CKMAGDIYKECANFLIVLKAHNTHLVACGAFHPVCTYVDVGHPEQDNVFLKDE
SFEENGKRSKPYDPLKLTASLILDEIYSGTADMGDPFAFRLTGNHPIRTQHD
SRMLNDPRTGAYIVPESDNDEDDKYVFFFRNADIDGHTGATHTARGLQCKNDPFG
HSLVKNWTFPLKARLCSVPGPNGIDTHDELQDVILNNSKDPNPVAVAVFTTSSN

IFKSAVCMYSLTDIRRVFLGPYAHARDGPNTQWPFQGRVYPREPCTPRTQFGFDS
TKDLDEVIWFAHSHAPAMNYPVPIINNPIIITKEVDYQFOIVDRLEADGQYDM
FIGTDMGTILKVVSVKETEWTDLBLEETWLESTPQPAISMEISTQOOLYIGSSV
GVSQLPDRCDVYGSCACCAECLARDPYCAMSCSSCFRFPKAPRTTRQDIRNGDPLT
HCSDLQHQDDPHRQSLERKITYCHNSSTPIECSSKSRALVFPQKQNEKDEIK
VDERKITHGILLTKRDSGLIYCAVHGFQTLKTLETITDHELDLHKE
DEGDSHKHKEPSNSMSPQKIWRDPMQLINHPNLTMDEFCQVWRDRKQROKN
GNVQVSNTKWHLQENKGRNRRRTIEFRABRSV"
BASE COUNT 744 a 444 c 517 g 620 t
ORIGIN
Query Match 0.8%; Score 23; DB 5; Length 2325;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 416 GCAAGGATCATATTTCATT 438
|||||
Db 217 GCAAGGATCATATTTCATT 239
RESULT 55
AC100365 51312 bp DNA linear HTG 22-NOV-2001
LOCUS Mus musculus clone RP23-128120, LOW-PASS SEQUENCE SAMPLING.
ACCESSION AC100365
VERSION AC100365.1 GI:17047731
KEYWORDS
HTG; HTGS PHASE0.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 51312)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Mus musculus, clone RP23-128120
Unpublished
2 (bases 1 to 51312)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choedel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., DeRellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ghinde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagde, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McKean, P., McKernan, K., McNeely, R., Meldrum, J.,
Meneus, L., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, D., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembo, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Web site: http://www-seq.wi.mit.edu
Contact: sequence submissions@genome.wi.mit.edu
----- Project Information
Center project name: L15170
Center clone name: 128_1_20

* NOTE: This record contains 65 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1 687: contig of 687 bp in length
688 787: gap of 100 bp
788 1479: contig of 692 bp in length
1480 1579: gap of 100 bp
1580 2256: contig of 677 bp in length
2257 2356: gap of 100 bp
2357 3050: contig of 694 bp in length
3051 3150: gap of 100 bp
3151 3830: contig of 680 bp in length
3831 3930: gap of 100 bp
3931 4612: contig of 682 bp in length
4613 4712: gap of 100 bp
4713 5374: contig of 662 bp in length
5375 5474: gap of 100 bp
5476 6159: contig of 685 bp in length
6160 6259: gap of 100 bp
6260 6979: contig of 720 bp in length
6980 7079: gap of 100 bp
7080 7753: contig of 674 bp in length
7754 7853: gap of 100 bp
7854 8520: contig of 667 bp in length
8521 9307: contig of 687 bp in length
8621 9407: gap of 100 bp
9308 10103: contig of 696 bp in length
9408 10203: gap of 100 bp
10104 10912: contig of 709 bp in length
10204 11012: gap of 100 bp
10913 11688: contig of 676 bp in length
11013 11788: gap of 100 bp
11689 12490: contig of 702 bp in length
11789 12590: gap of 100 bp
12491 13283: contig of 693 bp in length
12591 13383: gap of 100 bp
13284 14049: contig of 666 bp in length
13384 14149: gap of 100 bp
14050 14837: contig of 688 bp in length
14150 14937: gap of 100 bp
14838 15643: contig of 706 bp in length
14938 15743: gap of 100 bp
15644 16442: contig of 699 bp in length
15744 16542: gap of 100 bp
16443 17220: contig of 678 bp in length
16543 17320: gap of 100 bp
17221 18009: contig of 689 bp in length
17321 18109: gap of 100 bp
18010 18814: contig of 705 bp in length
18110 18914: gap of 100 bp
18815 19602: contig of 688 bp in length
18915 19702: gap of 100 bp
19603 20390: contig of 688 bp in length
19703 20490: gap of 100 bp
20391 21173: contig of 683 bp in length
20491 21273: gap of 100 bp
21174 21945: contig of 672 bp in length
21274 22045: gap of 100 bp
21946 22742: contig of 697 bp in length
22046 22842: gap of 100 bp
22743 23541: contig of 699 bp in length
22843 23641: gap of 100 bp
23542 24340: contig of 699 bp in length
23641 24440: gap of 100 bp
24341

24441 25125: contig of 685 bp in length
25126 25225: gap of 100 bp
25226 25937: contig of 712 bp in length
25938 26037: gap of 100 bp
26038 26733: contig of 695 bp in length
26733 26832: gap of 100 bp
26833 27517: contig of 685 bp in length
27518 27618: gap of 100 bp
27618 28286: contig of 669 bp in length
28287 28386: gap of 100 bp
28387 29103: contig of 717 bp in length
29104 29203: gap of 100 bp
29204 29886: contig of 683 bp in length
29887 30670: gap of 100 bp
29887 30670: contig of 684 bp in length
30671 30770: gap of 100 bp
30771 31493: contig of 723 bp in length
31494 31593: gap of 100 bp
31594 32274: contig of 680 bp in length
32274 33373: gap of 100 bp
33373 33078: contig of 705 bp in length
33079 33178: gap of 100 bp
33179 33877: contig of 693 bp in length
33878 33977: gap of 100 bp
33978 34669: contig of 692 bp in length
34670 34769: gap of 100 bp
34770 35415: contig of 645 bp in length
35415 35515: gap of 100 bp
35515 36222: contig of 708 bp in length
36223 36322: gap of 100 bp
36323 37007: contig of 685 bp in length
37008 37107: gap of 100 bp
37108 37830: contig of 723 bp in length
37831 37930: gap of 100 bp
37931 38632: contig of 702 bp in length
38633 38732: gap of 100 bp
38733 39392: contig of 660 bp in length
39393 39492: gap of 100 bp
39493 40155: contig of 663 bp in length
40156 40255: gap of 100 bp
40256 40963: contig of 708 bp in length
40964 41063: gap of 100 bp
41063 41763: contig of 700 bp in length
41764 41863: gap of 100 bp
41864 42581: contig of 718 bp in length
42582 42681: gap of 100 bp
42682 43385: contig of 704 bp in length
43386 43485: gap of 100 bp
43486 44201: contig of 716 bp in length
44202 44301: gap of 100 bp
44302 44978: contig of 677 bp in length
44979 45078: gap of 100 bp
45079 45771: contig of 693 bp in length
45772 45871: gap of 100 bp
45872 46574: contig of 703 bp in length
46575 46674: gap of 100 bp
46675 47371: contig of 697 bp in length
47372 47471: gap of 100 bp
47472 48164: contig of 693 bp in length
48165 48264: gap of 100 bp
48265 48951: contig of 687 bp in length
48952 49051: gap of 100 bp
49052 49722: contig of 671 bp in length
49723 49822: gap of 100 bp
49823 50517: contig of 695 bp in length
50518 51312: gap of 100 bp
51312 50618: contig of 695 bp in length.

FEATURES

1..51312 Mus musculus
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP23-128120"

BASE COUNT 12962 a 8905 c 9030 g 13847 t 6568 others
 ORIGIN /clone_lib="RPCT-23 Female Mouse BAC"

Query Match 0.8%; Score 23; DB 2; Length 51312;
 Best Local Similarity 100.0%; Pred.No.1.3;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 732 CATCCCTTTAATAGATGAGAA 754
 |||||
 DB 13056 CATCCCTTTAATAGATGAGAA 13078

RESULT 56
 AC120544/c
 LOCUS AC120544
 DEFINITION Mus musculus clone RP23-299N6, *** SEQUENCING IN PROGRESS ***, 3
 ORDERED pieces.
 AC120544
 AC120544.8 GI:31442497
 HTG: HTGS PHASE2; HTGS_FOLDTOP; HTGS_ACTIVEFIN.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Birren, B., Nussbaum, C. and Lander, E.
 Mus musculus, clone RP23-299N6
 Unpublished
 2 (bases 1 to 225372)
 Birren, B., Linton, L., Nussbaum, C., Lander, E., All, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
 Bouhagbeler, B., Brown, A., Camarata, J., Campoliano, A., Chang, J.,
 Chararo, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
 Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Ginde, S., Gord, S., Goyette, M., Graham, L., Grand, P., Hume, W.,
 Hagsos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kelle, C., Lacroque, K., Lamazares, R.,
 Landers, T., Lehoczkay, J., Levine, R., Lindblad-Toh, K., Liu, G.,
 Maclean, C., MacDonald, P., Major, J., Margulis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., Meltrin, J., Meneses, L.,
 Mihova, T., Mienna, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Piere, N., Pollara, V.,
 Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission

TITLE Submitted (08-MAY-2002) Whitehead Institute/MIT Center for Genome
 JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE 3 (bases 1 to 225372)
 AUTHORS Birren, B., Nussbaum, C., Lander, E., Abouelleil, A., Allen, N.,
 Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
 Boguslavsky, L., Bouhagbeler, B., Camarata, J., Chang, J., Choepel, Y.,
 Collymore, A., Cook, A., Cooke, P., Corum, B., Dearellano, K.,
 Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Fero, S.,
 Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
 Graham, L., Grand-Pierre, N., Hafez, N., Hagopian, D., Hagsos, B.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, G., Liu, A., Mabbitt, R., Maclean, C.,
 MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Meltrin, J., Meneses, L., Mihova, T., Mienna, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Piere, N.,
 Rachupka, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schnupack, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,

Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission

TITLE Submitted (06-JUN-2003) Whitehead Institute/MIT Center for Genome
 JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA
 REFERENCE On Jun 6, 2003 this sequence version replaced gi:31339910.
 COMMENT All repeats were identified using RepeatMasker:
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIRB
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 Project Information
 Center project name: 299_N6

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 3 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

1 39864: contig of 39864 bp in length
 * 39865 39864: gap of 100 bp
 * 39965 74622: contig of 34658 bp in length
 * 74623 74722: gap of 100 bp
 * 74723 225372: contig of 150650 bp in length.
 Location/Qualifiers

1..225372
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /clone="RP23-299N6"
 /clone_lib="RPCT-23 Female Mouse BAC"
 BASE COUNT 67480 a 43750 c 43731 g 69944 t 467 others
 ORIGIN

Query Match 0.8%; Score 23; DB 2; Length 225372;
 Best Local Similarity 100.0%; Pred.No.1.2;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1076 TGCTCAGTCCGACGTCCTCAATGG 1098
 |||||
 DB 179450 TGCTCAGTCCGACGTCCTCAATGG 179428

RESULT 57
 CGU28241
 LOCUS CGU28241 883 bp mRNA linear VRT 07-FEB-1996
 DEFINITION Gallus gallus collapsin-3 mRNA, partial cds.
 ACCESSION U28241
 VERSION U28241.1 GI:886810
 KEYWORDS
 SOURCE Gallus gallus (chicken)
 ORGANISM Gallus gallus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 883)
 Luo, Y., Shepherd, I., Li, J., Renzi, M.J., Chang, S. and Raper, J.A.
 A family of molecules related to collapsin in the embryonic chick
 nervous system
 JOURNAL Neuron 14 (6), 1131-1140 (1995)
 MEDLINE 95329269
 PUBMED 7605628
 REFERENCE 2 (bases 1 to 883)
 AUTHORS Luo, Y.

TITLE Direct Submission
JOURNAL Submitted (02-JUN-1995) Yuling Luo, Neuroscience, Univ.
Pennsylvania, 105 Johnson Pavilion, Philadelphia, PA 19104, USA

FEATURES
source
1.883
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
<1-->883
/codon_start=1
/product="collapsin-3"
/protein_id="AA86897.1"
/db_xref="GI:886811"
/translation="DDKIFPFEKRLTNSGSKQIHSMARICPNDTGQORSLVINK
TFLKARVCSVMDGDETEYDELEVDVLETDNPRITLVYGIPTSSSIFKSAVC
VYHLSDIOTVFNPGPAHKGSGNHOLIIPYRPGTCGCAFTPMPTTKRPDD
VTIRNDPLATNPIYPIHKRPLIRITGDTYKTIANDRNADGRCHVIFLTGDDG
TVQKRVVLPINFSASGELILELEVFQSNSPITTKISKQQLYVSSEGVTVPLH
RCRIYGTACACCLARDE"

BASE COUNT 268 a 185 c 195 g 235 t

ORIGIN

Query Match 0.8%; Score 22; DB 5; Length 883;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1778 TGTGGCTGGCCGACGACCTT 1799
|||||
Db 862 TGTGGCTGGCCGACGACCTT 883

RESULT 58
AX359932/c 5732 bp DNA linear PAT 13-FEB-2002

LOCUS AX359932
DEFINITION Sequence 6 from Patent WO0202783.
ACCESSION AX359932
VERSION AX359932.1 GI:18675571

KEYWORDS
SOURCE synthetic construct
ORGANISM artificial sequences.

REFERENCE
1
AUTHORS Bleck,G.T.
TITLE Expression vectors
JOURNAL Patent: WO 0202783-A 6 10-JAN-2002;
Gala Design, Inc. (US)
location/Qualifiers

FEATURES
source
1.5732
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"
/note="synthetic"

BASE COUNT 1390 a 1583 c 1425 g 1334 t

ORIGIN

Query Match 0.8%; Score 22; DB 6; Length 5732;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATGTT 22
|||||
Db 5112 AATCTTTATTTATCGATGTT 5091

RESULT 59
AX382146/c 5732 bp DNA linear PAT 18-MAR-2002

LOCUS AX382146
DEFINITION Sequence 6 from Patent WO0202738.
ACCESSION AX382146
VERSION AX382146.1 GI:19576955

KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct

artificial sequences.

REFERENCE
1
AUTHORS Bremel,R.D., Miller,L.U., Bleck,G.T. and York,D.
TITLE Host cells containing multiple integrating vectors
JOURNAL Patent: WO 0202738-A 6 10-JAN-2002;
Gala Design, Inc. (US)
location/Qualifiers

FEATURES
source
1.5732
/organism="synthetic construct"
/mol_type="genomic DNA"
/db_xref="taxon:32630"

BASE COUNT 1390 a 1583 c 1425 g 1334 t

ORIGIN

Query Match 0.8%; Score 22; DB 6; Length 5732;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATGTT 22
|||||
Db 5112 AATCTTTATTTATCGATGTT 5091

RESULT 60
A80476/c 7311 bp DNA circular PAT 21-JAN-2000

LOCUS A80476
DEFINITION Sequence 3 from Patent WO9943795.
ACCESSION A80476
VERSION A80476.1 GI:6731330

KEYWORDS
SOURCE unidentified
ORGANISM unidentified

REFERENCE
1 (bases 1 to 7311)
AUTHORS Benkel,B.F. and Falconer,M.M.
TITLE TRANS-SOMATICS WITH GENE TRANSFER INTO MAMMARY EPITHELIAL CELLS
JOURNAL Patent: WO 9943795-A 3 02-SEP-1999;
BENKEL BERNHARD F (CA); FALCONER MARCIA M (CA)
location/Qualifiers

FEATURES
source
1.7311
/organism="unidentified"
/mol_type="genomic DNA"
/db_xref="taxon:32644"

BASE COUNT 1745 a 1935 c 1843 g 1788 t

ORIGIN

Query Match 0.8%; Score 22; DB 6; Length 7311;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATGTT 22
|||||
Db 4330 AATCTTTATTTATCGATGTT 4309

RESULT 61
BD136840/c 7311 bp DNA linear PAT 18-SEP-2002

LOCUS BD136840
DEFINITION Transsomatrics associated with gene transfer into mammary epithelial cells.
ACCESSION BD136840
VERSION BD136840.1 GI:23231785
KEYWORDS JP 2002504365-A/3.
SOURCE unidentified
ORGANISM unidentified

REFERENCE
1 (bases 1 to 7311)
AUTHORS Gavora,J.S., Falconer,M.M., Nguyen,T.H. and Benkel,B.F.
TITLE Transsomatrics associated with gene transfer into mammary epithelial cells
JOURNAL Patent: JP 2002504365-A 3 12-FEB-2002;
HER MAJESTY IN RIGHT OF CANADA AS REPRESENTED BY THE MINISTER OF
AGRICULTURE AND AGRI FOOD CANADA

COMMENT OS Unidentified
 PN JP 2002504365-A/3
 PD 12-FEB-2002
 PR 25-JUN-1998 JP 2000533535
 PR 24-FEB-1998 CA 2224108
 PI JAN S GAYORA, MARCIA M FALCONER, TUY H NGUYEN, BERNHARD F BENKEL
 PC A01K67/027, C12N9/64, C12N15/09, C12P21/00, C12N15/00 CC
 Strandedness: Double;
 Topology: Circular;
 CC Transonetics associated with gene transfer into mammary CC
 epithelial cells
 FH Key Location/Qualifiers
 FT source 1..7311
 FT Location/Qualifiers
 1..7311
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

BASE COUNT 1745 a 1935 c 1843 g 1788 t
 ORIGIN

Query Match 0.8%; Score 22; DB 6; Length 7311;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AATCTTTATTTATCGATGTT 22
 |||||
 4330 AATCTTTATTTATCGATGTT 4309

RESULT 62
 AR302095/c 7352 bp DNA linear PAT 12-JUN-2003
 LOCUS Sequence 4 from patent US 6541197.
 DEFINITION AR302095
 ACCESSION AR302095.1 GI:31690129
 VERSION
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 7352)
 AUTHORS Link, C.J., Jr., Levy, J.P., Wang, S. and Seregina, T.
 TITLE Vehicles for stable transfer of green fluorescent protein gene and
 methods of use for same.
 JOURNAL Patent: US 6541197-A 4 01-APR-2003;
 FEATURES Location/Qualifiers
 source 1..7352
 /organism="unknown"
 BASE COUNT 1735 a 1997 c 1877 g 1743 t
 ORIGIN

Query Match 0.8%; Score 22; DB 6; Length 7352;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AATCTTTATTTATCGATGTT 22
 |||||
 4371 AATCTTTATTTATCGATGTT 4350

RESULT 63
 AR302092/c 7353 bp DNA linear PAT 12-JUN-2003
 LOCUS Sequence 1 from patent US 6541197.
 DEFINITION AR302092
 ACCESSION AR302092.1 GI:31690126
 VERSION
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unclassified.
 REFERENCE 1 (bases 1 to 7353)
 AUTHORS Link, C.J., Jr., Levy, J.P., Wang, S. and Seregina, T.

TITLE Vehicles for stable transfer of green fluorescent protein gene and
 methods of use for same
 JOURNAL Patent: US 6541197-A 1 01-APR-2003;
 FEATURES Location/Qualifiers
 source 1..7353
 /organism="unknown"
 BASE COUNT 1704 a 2062 c 1893 g 1694 t
 ORIGIN

Query Match 0.8%; Score 22; DB 6; Length 7353;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AATCTTTATTTATCGATGTT 22
 |||||
 4372 AATCTTTATTTATCGATGTT 4351

RESULT 64
 AX107929/c 7591 bp DNA linear PAT 30-APR-2001
 LOCUS Sequence 57 from Patent WO0125466.
 DEFINITION AX107929
 ACCESSION AX107929
 VERSION AX107929.1 GI:13923312
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 REFERENCE 1
 AUTHORS Slingsby, J., Kingsman, S.M., Rohll, J.O. and Slade, A.O.
 TITLE Producer cell for the production of retroviral vectors
 JOURNAL Patent: WO 0125466-A 57 12-APR-2001;
 Oxford Biomedica (UK) Limited (GB)
 FEATURES Location/Qualifiers
 source 1..7591
 /organism="synthetic construct"
 /mol_type="genomic DNA"
 /db_xref="taxon:32630"
 /note="Synthetic sequence"

BASE COUNT 1752 a 2116 c 2007 g 1726 t
 ORIGIN

Query Match 0.8%; Score 22; DB 6; Length 7591;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AATCTTTATTTATCGATGTT 22
 |||||
 3100 AATCTTTATTTATCGATGTT 3079

RESULT 65
 A80477/c 7885 bp DNA circular PAT 21-JAN-2000
 LOCUS Sequence 4 from Patent WO9943795.
 DEFINITION A80477
 ACCESSION A80477
 VERSION A80477.1 GI:6731331
 KEYWORDS
 SOURCE unidentified
 ORGANISM unidentified
 REFERENCE 1 (bases 1 to 7885)
 AUTHORS Benkel, B.F. and Falconer, M.M.
 TITLE TRANS-SOMATICS WITH GENE TRANSFER INTO MAMMARY EPITHELIAL CELLS
 JOURNAL Patent: WO 9943795-A 4 02-SEP-1999;
 BENKEL, BERNHARD F (CA); FALCONER, MARCIA M (CA)
 FEATURES Location/Qualifiers
 source 1..7885
 /organism="unidentified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

BASE COUNT 1876 a 2092 c 1996 g 1921 t
 ORIGIN

Query Match 0.8%; Score 22; DB 6; Length 7885;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AATCTTTATTTATCGATGTT 22
 |||||
 DB 4904 AATCTTTATTTATCGATGTT 4883

RESULT 66
 BD136841/c 7885 bp DNA linear PAT 18-SEP-2002
 LOCUS Transsomatics associated with gene transfer into mammary epithelial
 DEFINITION cells.
 ACCESSION BD136841 GI:23231786
 VERSION JP 2002504365-A/4.
 KEYWORDS
 SOURCE unclassified
 ORGANISM unclassified

REFERENCE 1 (bases 1 to 7885)
 Gavora, J.S., Falconer, M.M., Nguyen, T.H. and Benkel, B.F.
 Transsomatics associated with gene transfer into mammary epithelial
 TITLE cells
 JOURNAL Patent: JP 2002504365-A 4 12-FEB-2002;
 HER MAJESTY IN RIGHT OF CANADA AS REPRESENTED BY THE MINISTER OF
 AGRICULTURE AND AGRI FOOD CANADA

COMMENT OS Unidentified
 PN JP 2002504365-A/4
 PD 12-FEB-2002
 PF 25-JUN-1998 JP 2000533535
 PR 24-FEB-1998 CA 2224108
 PI JAN S GAVORA, MARCIA M FALCONER, THUY H NGUYEN, BERNHARD F BENKEL
 PC A01K67/027,C12N9/64,C12N15/09,C12P21/00,C12N15/00 CC
 Strandedness: Double;
 CC Topology: Circular;
 CC Transsomatics associated with gene transfer into mammary CC
 FH Key location/Qualifiers
 FT source 1..7885
 ORIGIN Location/Qualifiers
 1..7885
 /organism="unclassified"
 /mol_type="genomic DNA"
 /db_xref="taxon:32644"

BASE COUNT 1876 a 2092 c 1996 g 1921 t
 ORIGIN

Query Match 0.8%; Score 22; DB 6; Length 7885;
 Best Local Similarity 100.0%; Pred. No. 5.3;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AATCTTTATTTATCGATGTT 22
 |||||
 DB 4904 AATCTTTATTTATCGATGTT 4883

RESULT 67
 AX359933/c 9183 bp DNA linear PAT 13-FEB-2002
 LOCUS Sequence 7 from Patent WO0202783.
 DEFINITION
 ACCESSION AX359933
 VERSION AX359933.1 GI:18675572.
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 TITLE artificial sequences.
 REFERENCE 1
 AUTHORS Bleck, G.T.
 TITLE Expression vectors
 JOURNAL Patent: WO 0202783-A 7 10-JAN-2002;

FEATURES
 source
 Location/Qualifiers
 1..9183
 /organism="synthetic construct"
 /mol_type="genomic DNA"
 /db_xref="taxon:32630"
 /note="Synthetic"

BASE COUNT 2133 a 2498 c 2399 g 2153 t
 ORIGIN

Query Match 0.8%; Score 22; DB 6; Length 9183;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AATCTTTATTTATCGATGTT 22
 |||||
 DB 8563 AATCTTTATTTATCGATGTT 8542

RESULT 68
 AX382147/c 9183 bp DNA linear PAT 18-MAR-2002
 LOCUS Sequence 7 from Patent WO0202738.
 DEFINITION
 ACCESSION AX382147
 VERSION AX382147.1 GI:19576956
 KEYWORDS
 SOURCE synthetic construct
 ORGANISM synthetic construct
 TITLE artificial sequences.

REFERENCE 1
 BREMEL, R.D., MILLER, L.U., BLECK, G.T. and YORK, D.
 Host cells containing multiple integrating vectors
 TITLE Patent: WO 0202738-A 7 10-JAN-2002;
 JOURNAL Gala Design, Inc. (US)
 FEATURES
 source
 Location/Qualifiers
 1..9183
 /organism="synthetic construct"
 /mol_type="genomic DNA"
 /db_xref="taxon:32630"

BASE COUNT 2133 a 2498 c 2399 g 2153 t
 ORIGIN

Query Match 0.8%; Score 22; DB 6; Length 9183;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AATCTTTATTTATCGATGTT 22
 |||||
 DB 8563 AATCTTTATTTATCGATGTT 8542

RESULT 69
 AC006525/c 90604 bp DNA linear PRI 19-APR-2002
 LOCUS Homo sapiens chromosome 5 clone RP1-401.7, complete sequence.
 DEFINITION
 ACCESSION AC006525
 VERSION AC006525.1 GI:4225899
 KEYWORDS
 SOURCE HTG.
 ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 90604)
 KIMMERLY, W., BOND, M., CHENG, J., CONNOLLY, K.S., KADNER, K.,
 MIGUEL, T., MILLER, C., PITLUCK, S., POLLARD, M., ROJESKI, H.,
 SUBRAMANIAN, S., WHEELAND, A. and MARTIN, C.H.
 Sequencing of human chromosome 5
 TITLE Unpublished
 REFERENCE 2 (bases 1 to 90604)
 RIECKE, D.O.
 TITLE Large Scale Sequence Analysis and Annotation with the Sequence
 JOURNAL Comparison Analysis (SCAN) System
 Unpublished

REFERENCE 3 (bases 1 to 90604)
 AUTHORS Kimerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Kadner,K.,
 Miguel,T., Miller,C., Pitluck,S., Pollard,M., Rojsek,I.H.,
 Subramanian,S., Wheeland,A. and Martin,C.H.
 TITLE Direct Submission
 JOURNAL Submitted (05-FEB-1999) Human Genome Center, DOE Joint Genome
 Institute, Lawrence Berkeley National Laboratory, MS 74-157,
 Berkeley, CA 94720, U.S.A.
 REFERENCE 4 (bases 1 to 90604)
 AUTHORS Muntz,M.O.
 TITLE Direct Submission
 JOURNAL Submitted (19-APR-2002) Biosciences, Joint Genome Institute/ Los
 Alamos National Laboratory, MS M888, Los Alamos, New Mexico 87545,
 U.S.
 COMMENT Sequence submitted by:
 DOE Joint Genome Institute.
 FEATURES
 source
 1. 90604
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="RP1-4017"
 147. .264
 /rpt_family="MIR"
 871. .1157
 /rpt_family="Alu"
 complement(2190. .8204)
 /rpt_family="LI"
 complement(6746. .6964)
 /rpt_family="MER25"
 8436. .8715
 /rpt_family="Alu"
 9507. .9715
 /rpt_family="MLT"
 complement(11635. .11800)
 /rpt_family="LI"
 18197. .18493
 /rpt_family="Alu"
 complement(18701. .19051)
 /rpt_family="THE1"
 complement(19056. .19622)
 /rpt_family="THE"
 complement(19073. .20447)
 /rpt_family="MSTAR"
 complement(19655. .20251)
 /rpt_family="THR"
 complement(20961. .21052)
 /note="GRAIL 2 excellent exon, frame 0"
 22417. .22785
 /rpt_family="THE1"
 complement(24301. .25358)
 /rpt_family="LI"
 complement(26046. .26221)
 /note="GRAIL 2 excellent exon, frame 1"
 complement(26626. .26830)
 /rpt_family="LI"
 complement(30565. .30815)
 /rpt_family="MER44C"
 complement(32234. .33083)
 /rpt_family="LI"
 complement(33304. .35878)
 /rpt_family="LI"
 complement(35886. .36070)
 /rpt_family="Alu"
 36037. .36270
 /rpt_family="Alu"
 complement(36268. .42451)
 /rpt_family="LI"
 complement(40878. .41365)
 /rpt_family="MER25"
 complement(42728. .42898)
 /rpt_family="Alu"

repeat_region complement(42916. .43259)
 /rpt_family="Alu"
 repeat_region 44045. .44230
 /rpt_family="MER3"
 repeat_region 44096. .44142
 /rpt_family="MER33"
 repeat_region complement(45460. .45968)
 /rpt_family="LI"
 repeat_region complement(52284. .52407)
 /rpt_family="MIR"
 repeat_region complement(52989. .53321)
 /rpt_family="Alu"
 repeat_region complement(53942. .54100)
 /rpt_family="LI"
 repeat_region 55602. .55886
 /rpt_family="Alu"
 repeat_region complement(56037. .56502)
 /rpt_family="LI"
 repeat_region 59874. .60193
 /rpt_family="Alu"
 repeat_region 60477. .60752
 /rpt_family="Alu"
 repeat_region 60820. .61300
 /rpt_family="LI"
 repeat_region 61348. .61619
 /rpt_family="LI"
 repeat_region 62747. .63044
 /rpt_family="Alu"
 misc_feature complement(63623. .63754)
 /note="GRAIL 2 excellent exon, frame 0"
 64221. .64406
 /rpt_family="MER5"
 repeat_region complement(64263. .64381)
 /rpt_family="MER5"
 repeat_region complement(66455. .66792)
 /rpt_family="MER"
 repeat_region complement(67054. .67310)
 /rpt_family="Alu"
 repeat_region complement(67511. .69174)
 /rpt_family="LI"
 repeat_region complement(69265. .69379)
 /rpt_family="MIR"
 70504. .73086
 /rpt_family="LI"
 repeat_region 71805. .72106
 /rpt_family="Alu"
 misc_feature 72865. .72937
 /note="GRAIL 2 excellent exon, frame 2"
 repeat_region complement(73156. .74083)
 /rpt_family="LI"
 76196. .76310
 /note="GRAIL 2 excellent exon, frame 2"
 76342. .76625
 /rpt_family="Alu"
 repeat_region complement(77911. .78203)
 /rpt_family="THE1"
 repeat_region complement(78204. .79756)
 /rpt_family="MSTAR"
 repeat_region complement(79761. .80107)
 /rpt_family="THE1"
 repeat_region complement(83286. .83483)
 /rpt_family="LI"
 repeat_region complement(84022. .84091)
 /rpt_family="LI"
 repeat_region complement(84857. .84976)
 /rpt_family="LI"
 repeat_region 85589. .85681
 /rpt_family="LTR"
 repeat_region complement(86063. .86337)
 /rpt_family="Alu"
 repeat_region 86485. .86693
 /rpt_family="LI"
 repeat_region complement(87822. .90604)

```

BASE COUNT      /rpl family="11"
ORIGIN          25086 a 16222 c 16894 g 32402 t

Query Match      0.8%; Score 22; DB 9; Length 90604;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 287 CCAGGCTGAATATATCTACA 308
Db 10922 CCAGGCTGAATATATCTACA 10901

RESULT 70
AC112426.0
WPCOMMENT
Sequence split into 4 fragments LOCUS AC112426 Accession AC112426
Fragment Name Begin End
AC112426_0 1 110000
AC112426_1 100001 210000
AC112426_2 200001 310000
AC112426_3 300001 374204
LOCUS AC112426 374204 bp DNA linear HTG 08-OCT-2002
DEFINITION Rattus norvegicus clone CH230-312K14, *** SEQUENCING IN PROGRESS
AC112426
AC112426.4 GI:23270041
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
KEYWORDS Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 374204)
Muzny,D,Marie, Metzker,M,lee, Abramson,S, Adams,C, Alder,J,
Allen,C, Allen,H, Alsbrooks,S, Amin,A, Anguiano,D,
Ayalebechi,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H,
Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F,
Biswal,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M,
Bryant,N, Buhay,C, Burch,P, Butrell,K, Calderon,E,
Cardenas,J, Carter,K, Cavazos,I, Cesari,H, Center,A,
Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J,
Claveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L,
Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D,
Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K,
Drepper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K,
Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G,
Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P,
Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M,
Geregeorgis,B, Geer,K, Gill,R, Grady,M, Guerra,W, Guayra,W,
Guevarra,P, Haaland,W, Hamil,C, Hamilton,C, Hamilton,K,
Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J,
Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hogues,M,
Hollins,B, Howells,S, Hulik,S, Hume,J, Idlebird,D, Jackson,A,
Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolivet,A,
Karpacky,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C,
Kovis,C, Kraft,C,L, Lebow,H, Lavan,J, Lewis,L, Li,Z, Liu,J,
Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J,
Lorenshewa,L, Louisged,H, Lozano,R,J, Lu,X, Ma,J,
Maheshwari,M, Mahindartine,M, Mahmoud,M, Malloy,K, Mangum,A,
Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E,
Mathany,S, McLeod,M,P, McNeill,T,Z, Meenen,E,
Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S,
Morgan,M, Morris,K, Morris,S, Munidasa,M, Murphy,M, Nair,L,
Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S,
Nwokilemeh,O, Okwomou,G, Olarunpungoon,A, Pal,S, Parks,K,
Patternak,S, Paul,H, Perez,A, Perez,L, Pfannkoch,C,
Plommer,F, Poindexter,A, Popovic,D, Primus,B, Pu,L,L,
Puzo,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M,A, Reigh,R,
Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F,
Rivers,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J,
Sanders,W, Savery,G, Scherer,S, Scott,G, Shatman,S, Shen,H,
Shetty,J, Shvartsbeyn,A, Sleson,I, Sitter,C,D, Smaj,D,

```

```

TITLE
JOURNAL
REFERENCES
AUTHORS
TITLE
JOURNAL
REFERENCES
AUTHORS
TITLE
JOURNAL
COMMENT
Submitted (08-OCT-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 23, 2002 this sequence version replaced gi:21743438.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/atlas/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GRTR
Center clone name: CH230-312K14
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 240799 bases at least Q40
Consensus quality: 246749 bases at least Q30
Consensus quality: 250548 bases at least Q20
Estimated insert size: 345761; sum-of-contigs estimation
Quality coverage: 3x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This sequence may represent more than one clone
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 222049: contig of 222049 bp in length
* 222050 222149: gap of unknown length
* 222150 236912: contig of 14763 bp in length
* 236913 237012: gap of unknown length
* 237013 301105: contig of 64093 bp in length
* 301106 301205: gap of unknown length
* 301206 305403: contig of 4198 bp in length
* 305404 305503: gap of unknown length
* 305504 352310: contig of 46807 bp in length

```

```

* 352311 352410: gap of unknown length
* 352411 353860: contig of 1450 bp in length
* 353861 353960: gap of unknown length
* 353961 355587: contig of 1627 bp in length
* 355588 355687: gap of unknown length
* 355688 357943: contig of 2256 bp in length
* 357944 358043: gap of unknown length
* 358044 360096: contig of 2053 bp in length
* 360097 360196: gap of unknown length
* 360197 361828: contig of 1632 bp in length
* 361829 361928: gap of unknown length
* 361929 364193: contig of 2265 bp in length
* 364194 364294: gap of unknown length
* 364294 364394: contig of 9911 bp in length.

```

```

FEATURES
  source
    1. 374204
      /organism="Rattus norvegicus"
      /mol_type="genomic DNA"
      /db_xref="taxon:10116"
      /clone="CH230-312K14"
      /cfe="CH230-312K14"
      /note="wgs contig"
  misc_feature
    220130..222049
      /note="wgs contig"
  misc_feature
    301206..303668
      /note="wgs contig"
  misc_feature
    303719..305403
      /note="wgs contig"
  misc_feature
    72927 a 53729 c 54751 g 71203 c 121594 others
  BASE COUNT
  ORIGIN

```

```

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 110000;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 2248 AGATGCTTAAGACCAAGAA 2269
Db 91867 AGATGCTTAAGACCAAGAA 91888

Sequence split into 4 fragments LOCUS AC112426 Accession AC112426
Fragment Name Begin End
AC112426_1 1 110000
AC112426_2 100001 210000
AC112426_3 200001 310000
AC112426_4 300001 374204
Continuation (3 of 4) of AC112426 from base 200001 (AC112426 Rattus norvegicus clone CH2

```

```

Query Match
Best Local Similarity 100.0%; Score 22; DB 2; Length 110000;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 2248 AGATGCTTAAGACCAAGAA 2269
Db 37340 AGATGCTTAAGACCAAGAA 37361

```

```

RESULT 72
LOCUS AC129994
DEFINITION Rattus norvegicus clone CH230-510F7, WORKING DRAFT SEQUENCE, 4
ACCESSION AC129994
VERSION AC129994.3 GI:25007674
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

```

REFERENCE

- 1 (bases 1 to 150638)
- Munzky, D., Marie, Metker, M., Lee, A., Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Albrooks, S., Amin, A., Angiano, D., Anyalebech, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, B., Benham, F., Biwalto, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cessari, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., p Souza, L., Davila, M., Davis, C., Day-Carroll, L., De Ande, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dint, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durkin, K., Duval, B., Evans, K., Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorjis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jallvet, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Kapathy, S., Kelly, S., Khan, Z., King, L., Koyar, C., Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhewa, L., Louiseged, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M. P., McNeill, T. Z., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundaasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackemeah, O., Okumu, G., Olampunagoo, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C., Plopper, F., Poindecker, A., Popovic, D., Primus, E., Pu, L., Puzo, M., Quiroz, J., Rachlin, B., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J., Sanders, W., Saverly, G., Scherer, S., Scott, G., Shatman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smaj, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorrell, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabot, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleciyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, J., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G., and Gibbs, R. A.
- Unpublished
- Direct Submission
- 2 (bases 1 to 150638)
- Worley, K. C.
- Direct Submission
- Submitted (08-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
- 3 (bases 1 to 150638)
- Rat Genome Sequencing Consortium.
- Direct Submission
- Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
- On Nov 15, 2002 this sequence version replaced gi:23196240.
- The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.bgsc.bcm.tmc.edu/projects/atl/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence

contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

-----Genome Center

Center: Baylor College of Medicine
Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

-----Project Information

Center project name: XAGT
Center clone name: CH230-510F7

-----Summary Statistics

Assembly program: Phrap; version 0.990329
Consensus quality: 145002 bases at least Q40
Consensus quality: 146043 bases at least Q30
Consensus quality: 146587 bases at least Q20
Estimated insert size: 148398; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 128673: contig of 128673 bp in length
* 128674 128773: gap of unknown length
* 128774 147523: contig of 18750 bp in length
* 147524 147623: gap of unknown length
* 147624 148701: contig of 1078 bp in length
* 148702 148801: gap of unknown length
* 148802 150638: contig of 1837 bp in length.

FEATURES

source

1. 150638
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-510F7"

misc_feature

2017..105883
/note="clone boundary
clone_end:Sp6
site:
end_sequence:EG117461"
6382..7103
/note="clone boundary
clone_end:T7
site:
end_sequence:EG117458"
128774..131862
/note="wgs_end-extension
clone_end:T7"

misc_feature

end_sequence:EG117461"
6382..7103
/note="clone boundary
clone_end:T7
site:
end_sequence:EG117458"
128774..131862
/note="wgs_end-extension
clone_end:T7"

misc_feature

BASE COUNT 38301 a 36249 c 36133 g 36407 t 3548 others
ORIGIN

Query Match 0.8%; Score 22; DB 2; Length 150638;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

GY 2461 GAAGCTGAAACAGAGAGACC 2482
|||||

DB 40496 GAAGCTGAAACAGAGAGACC 40517
|||||

RESULT 73

LOCUS

AC122712 166965 bp DNA linear HTG 25-MAY-2002
DEFINITION Homo sapiens chromosome 5 clone RP11-35J11, WORKING DRAFT SEQUENCE,
19 unordered pieces.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AC122712
AC122712.1 GI:21206275
HTG; HTGS PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 166965)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 5
Unpublished
2 (bases 1 to 166965)
DOE Joint Genome Institute.
Direct Submission
Submitted (25-MAY-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 420032
Center clone name: RP11-35J11

Summary Statistics

Consensus quality: 149343 bases at least Q40
Consensus quality: 156729 bases at least Q30
Consensus quality: 160416 bases at least Q20

Estimated insert size: 175000; agarose-fp estimation
Estimated insert size: 165165; sum-of-contigs estimation
Quality coverage: 12.41 in Q20 bases; agarose-fp estimation
Quality coverage: 13.14 in Q20 bases; sum-of-contigs estimation.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 1095: contig of 1095 bp in length
* 1096 1195: gap of unknown length
* 1196 2331: contig of 1136 bp in length
* 2332 2431: gap of unknown length
* 2432 3885: contig of 1454 bp in length
* 3886 3985: gap of unknown length
* 3986 5474: contig of 1489 bp in length
* 5475 5574: gap of unknown length
* 5575 6785: contig of 1211 bp in length
* 6786 6885: gap of unknown length
* 6886 8468: contig of 1583 bp in length
* 8469 8568: gap of unknown length
* 8569 11081: contig of 2513 bp in length
* 11082 11181: gap of unknown length
* 11182 14532: contig of 3351 bp in length
* 14533 14632: gap of unknown length
* 14634 18334: contig of 3702 bp in length
* 18335 18434: gap of unknown length
* 18435 28070: contig of 9636 bp in length
* 28071 28171: gap of unknown length
* 28172 34825: contig of 6655 bp in length
* 34826 34925: gap of unknown length
* 34926 43325: contig of 8400 bp in length
* 43326 43425: gap of unknown length
* 43426 56689: contig of 13264 bp in length
* 56690 56789: gap of unknown length
* 56790 70875: contig of 14086 bp in length
* 70876 70975: gap of unknown length
* 70976 70976: gap of unknown length
* 70977 86880: contig of 15905 bp in length
* 86881 86980: gap of unknown length
* 86981 106763: contig of 19783 bp in length
* 106764 106863: gap of unknown length

FEATURES

source

1. 216124
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-1N4"

BASE COUNT 41454 a 31758 c 30503 g 41391 t 71018 others

ORIGIN

Query Match 0.8%; Score 22; DB 2; Length 216124;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 22; Mismatch 0; Mismatches 0; Indels 0; Gaps 0;

CY 2248 AGATGGCTTAAGACCAAGAA 2269
189397 AGATGGCTTAAGACCAAGAA 189376

RESULT 75
AC106106/c

LOCUS AC106106 222086 bp DNA linear HTG 10-MAY-2003

DEFINITION Rattus norvegicus clone CH230-14317, WORKING DRAFT SEQUENCE.

ACCESSION AC106106

VERSION AC106106.5 GI:30521604

KEYWORDS HTG_PHASE2; HTG_DRAFT; HTG_FULLTOP.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus; Rattus

REFERENCE 1 (bases 1 to 222086)
Munzy, D. Marie, Metzker, M. Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biewald, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Bunay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, D., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, A., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guetara, W., Gunatane, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A., Jaspard, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lounsbury, L., Louised, H., Lozdo, R. D., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangun, A., Mangun, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mauney, S., McLeod, M. P., McNeill, T. Z., Meenen, B., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundaasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokediem, O., Okwom, G., Olarnpungoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poindexter, A., Popovic, D., Prims, E., Pu, L. L., Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rivers, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,

FEATURES

source

1. 222086
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-14317"

REFERENCE

TITLE JOURNAL

REFERENCE

AUTHORS

TITLE JOURNAL

COMMENT

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GHKJ
Center clone name: CH230-14317
Assembly Statistics
Consensus quality: 213592 bases at least Q40
Consensus quality: 216173 bases at least Q30
Consensus quality: 217783 bases at least Q20
Estimated insert size: 225854; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 222086: contig of 222086 bp in length.
Location/Qualifiers
1. 222086
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-14317"

Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Sivartsbeyn, A., Sisson, I., Sitter, C. D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sotter, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Swatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Wilson, R., Wleczky, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhou, D., von Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstock, G. and Gibbs, R. A.

Unpublished
2 (bases 1 to 222086)
Worley, K. C.
Submitted (12-UN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 222086)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:25094802.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome sequence only contigs will be indicated in the feature table.

misc_feature 1.3026
/note="wgs_config"
BASE COUNT 55222 a 54777 c 53885 g 55298 t 2904 others
ORIGIN

Query Match 0.8%; Score 22; DB 2; Length 222086;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2461 GAAAGGTAGAAACAGGAGACC 2482
|||
Db 26127 GAAAGGTAGAAACAGGAGACC 26106

Search completed: July 31, 2003, 22:28:36
Job time : 9595 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 19:21:21 ; Search time 5030 Seconds
(without alignments)
13089.627 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709
Sequence: 1 aatcttatttaccatg.....agcttttttccataacc 2709

Scoring table: OLIGO_NUC
Gapop 60.0, Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size: 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database:

EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_ges_hum:*
18: em_ges_inv:*
19: em_ges_pln:*
20: em_ges_vrt:*
21: em_ges_fun:*
22: em_ges_mam:*
23: em_ges_mus:*
24: em_ges_pro:*
25: em_ges_rtd:*
26: em_ges_dhg:*
27: em_ges_vrl:*
28: gb_gesl:*
29: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	723	26.7	723	12	B1870437
2	526	19.4	679	12	BM711125
3	418	15.4	702	12	BM723807
4	408	15.1	510	14	CD216784

5	391	14.4	889	10	BF667677
6	387	14.3	497	14	CA395261
7	382	14.1	382	13	BX283116
8	324	12.0	721	2	HSW084323
9	315	11.6	670	12	BM670627
10	286	10.6	839	10	BF667143
11	284	10.5	747	12	BQ018643
12	274	10.1	884	10	BF700780
13	265	9.8	532	2	HSW084322
14	260	9.6	695	3	BQ447821
15	218	8.0	355	10	BF742853
16	213	7.9	231	14	T29595
17	203	7.5	328	9	A1671991
18	196	7.2	910	14	R19784
19	194	7.2	214	9	AL529332
20	135	5.0	325	10	BES52070
21	120	4.8	430	9	AA424316
22	104	3.8	896	13	B0186364
23	101	3.7	570	10	BE033028
24	68	2.5	1005	12	BM801410
25	67	2.5	677	29	AG157495
26	59	2.2	604	12	B183160
27	57	2.1	598	13	BX282800
28	56	2.1	576	28	AQ533647
29	52	1.9	559	14	CA872840
30	51	1.9	394	9	AA716143
31	50	1.8	773	28	BH265538
32	49	1.8	124	10	BG315503
33	47	1.7	371	13	BY021039
34	47	1.7	401	13	BY018262
35	47	1.7	450	13	BY261117
36	47	1.7	470	12	B1465498
37	47	1.7	639	10	BB655104
38	47	1.7	653	10	BB663470
39	47	1.7	635	10	BB663737
40	47	1.7	657	10	BB658393
41	47	1.7	677	14	BY734980
42	47	1.7	731	13	B0708579
43	47	1.7	787	14	CB518391
44	47	1.7	2596	11	AK053115
45	41	1.5	318	13	BY303092
46	41	1.5	807	14	CB595903
47	41	1.5	900	14	CB196368
48	41	1.5	917	13	B0849424
49	38	1.4	394	14	CB776561
50	37	1.4	657	13	BQ779621
51	37	1.4	658	13	B0758296
52	37	1.4	660	10	BB612039
53	35	1.3	699	14	CA327319
54	35	1.3	773	13	B0515916
55	35	1.3	775	13	B0703431
56	34	1.3	374	9	AA861394
57	30	1.1	600	12	B1989239
58	30	1.1	662	10	BB628362
59	30	1.1	690	14	W98303
60	30	1.1	712	9	A1894009
61	29	1.1	291	10	BB377317
62	29	1.1	292	10	BB366496
63	29	1.1	729	14	CD349263
64	27	1.0	446	10	BB859723
65	27	1.0	539	9	AL918018
66	25	0.9	685	12	BG926854
67	25	0.9	555	14	CD203422
68	24	0.9	598	14	CD203073
69	24	0.9	614	14	CA502146
70	24	0.9	792	12	BG781873
71	24	0.9	825	13	B0255423
72	24	0.9	851	13	B0326664
73	24	0.9	897	9	B0267223
74	23	0.8	184	9	AM080828
75	23	0.8	401	12	BG898214
76	23	0.8	591	14	CA779226
77	23	0.8	608	13	BQ907117

BE667677	602121773
CA395261	CS62C03.Y
BX283116	BX283116
BX510254	Homo sapi
BM670627	UI-E-DX1-
BF667143	602121928
BQ018643	UI-H-DH1-
BF700780	602128585
Bx491494	Homo sapi
BQ447821	UI-H-EU1-
BT742853	IL2-BT073
T29595	EST86074 Hu
A1671991	wb72C04.x
R19784	YG28C12.r1
AL529332	AL529332
BES52070	hy03d11.x
AA424316	zv90D08.Y
B0186364	AGENCOURT
BE033028	133242 MA
BM801410	AGENCOURT
AG157495	Par trogl
B183160	UNI-P-FN-
BX282800	BX282800
AQ533647	RPCT-11-3
CA872840	K0920H08-
AA716143	zg6D02.s
BH265538	CH230-43J
BG315503	PO3.0.131
BY021039	BY021039
BY018262	BY018262
BY261117	BY261117
B1465498	1e18d11.Y
BB655104	BB655104
BB663470	BB663470
BB663737	BB663737
BB658393	BB658393
BY734980	BY734980
B0708579	UI-M-F10-
CB518391	UI-M-GH0-
AK053115	Mus muscu
BY303092	BY303092
CB595903	AGENCOURT
CB196368	AGENCOURT
B0849424	AGENCOURT
CB776561	AMGNIC.S
BQ779621	UI-R-FP0-
B0758296	UI-R-FP0-
BB612039	BB612039
CA327319	UI-M-FY0-
B0515916	AGENCOURT
B0703431	UI-M-F00-
AA861394	AK37C08.S
B1989239	4032-91 M
BB628362	BB628362
W98303	mg13h01.r1
A1894009	mg13h01.Y
BB377317	BB377317
BB366496	BB366496
CD349263	UI-M-FY0-
BB859723	BB859723
AL918018	AL918018
BG926854	HNC20-1-B
CD203422	LG_AMI_08
CD203073	LG_AMI_02
CA502146	WHEH043.C
BG781873	SEADMC001
B0255423	603414282
B0326664	603493391
B0267223	603815106
AM080828	xc38a05.x
BG898214	HOA38-1-B
CA779226	MDP384_8
BQ907117	N003D03.O

78	23	0.8	625	12	BG897406	BG897406 HOA13-1-F
79	23	0.8	638	12	BG926247	BG926247 HNC50-1-G
80	23	0.8	730	12	BG975571	BG975571 602845241
81	23	0.8	822	12	BI661527	BI661527 603305839
82	22	0.8	160	10	BF984015	BF984015 602306942
83	22	0.8	213	10	BE609913	BE609913 sq46508.Y
84	22	0.8	225	10	BB368206	BB368206 BB368206
85	22	0.8	330	9	AU036240	AU036240 AU036240
86	22	0.8	366	13	BG696984	BG696984 LL213001
87	22	0.8	408	12	BG897335	BG897335 HOA12-1-F
88	22	0.8	409	12	BG926168	BG926168 HNC2-1-D3
89	22	0.8	410	12	BG924214	BG924214 HNC25-1-G
90	22	0.8	415	10	BG025774	BG025774 602274765
91	22	0.8	418	12	BG897329	BG897329 HOA12-1-E
92	22	0.8	418	12	BG898828	BG898828 HOA57-1-C
93	22	0.8	435	12	BI813570	BI813570 KO14F09.O
94	22	0.8	441	14	CA871124	CA871124 KO908A03-C
95	22	0.8	447	12	BG901133	BG901133 HOA54-1-B
96	22	0.8	452	12	BG926300	BG926300 HNC3-1-E8
97	22	0.8	464	14	CA779986	CA779986 MPL384.3
98	22	0.8	474	12	BG927549	BG927549 HNC43-1-B
99	22	0.8	492	10	BF162689	BF162689 601769344
100	22	0.8	493	14	CB725110	CB725110 AMGNNUC.N

ALIGNMENTS

RESULT 1
BI870437
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BI870437 723 bp mRNA linear EST 11-OCT-2001
603395690F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405519 5',
mRNA sequence.
BI870437
BI870437.1 GI:16044110
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 723)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ARCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.jnl.gov
Plate: LHAM2034 row: C column: 24
High quality sequence stop: 723.
Location/Qualifiers
1: 723
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5405519"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_1b="NIH MGC 90"
/note="Organ: Liver; Vector: pCMV-SORTe; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC library."

FEATURES
source
BASE COUNT
ORIGIN
Query Match %

225 a 148 c 159 g 191 t
26.7%; Score 723; DB 12; Length 723;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 723; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	454	CAAGATTTTCAAAAGATTTGTGGCCAGTATCTTACACCAAGAGATGATGCAAGG	513
Db	1	CAAGATTTTCAAAAGATTTGTGGCCAGTATCTTACACCAAGAGATGCAAGG	60
Qy	514	GGCTGGAAGAAGATCTCTGAAGAATGTCTAATTTTCATCAAGTAAAGGCAATATA	573
Db	61	GGCTGGAAGAAGATCTCTGAAGAATGTCTAATTTTCATCAAGTAAAGGCAATATA	120
Qy	574	TGAGACTCACTTGTAGCGCTGTGGAACGGGGCTTTTCATCAATTTGCACTTATGA	633
Db	121	TGAGACTCACTTGTAGCGCTGTGGAACGGGGCTTTTCATCAATTTGCACTTATGA	180
Qy	634	AATTTGACATCATCTTGTAGGAGCAATATTTTAACTGAGAACTCAATTTGAAACGG	693
Db	181	AATTTGACATCATCTTGTAGGAGCAATATTTTAACTGAGAACTCAATTTGAAACGG	240
Qy	694	CCGTGGGAAGATTCATATGACCTTAAGCTGCTGACAGCATCCCTTTTAAATGAGAGA	753
Db	241	CCGTGGGAAGATTCATATGACCTTAAGCTGCTGACAGCATCCCTTTTAAATGAGAGA	300
Qy	754	ATTATCTCTGAACTGACGTGATTTTATGGGGGAGACTTTGCTATCTTCCGAATCT	813
Db	301	ATTATCTCTGAACTGACGTGATTTTATGGGGGAGACTTTGCTATCTTCCGAATCT	360
Qy	814	TGGGACACCAACCAATAGAGAGAGAGATATTCAGGTGGCTCAATGATCCAA	873
Db	361	TGGGACACCAACCAATAGAGAGAGATATTCAGGTGGCTCAATGATCCAA	420
Qy	874	GTTTCATTTAGTGGCCACCTCATCTCAGAGAGTACATCTGAAAGATGACAAAGTACTT	933
Db	421	GTTTCATTTAGTGGCCACCTCATCTCAGAGAGTACATCTGAAAGATGACAAAGTACTT	480
Qy	934	TTTCTTCCGTGAATGATGAGAGACATCTTGAAAAGCTACTACGCTTGAAT	993
Db	481	TTTCTTCCGTGAATGATGAGAGACATCTTGAAAAGCTACTACGCTTGAAT	540
Qy	994	AGGTGATATTCAGAAAGATCTTTGAGGGGACAGAGTCTGGTAATTAATGAGCAAC	1053
Db	541	AGGTGATATTCAGAAAGATCTTTGAGGGGACAGAGTCTGGTAATTAATGAGCAAC	600
Qy	1054	ATTCTCAAGCTCGTGTGATTTGCTCAGTGCAGGTCCCAATGGCATTCATCTAT	1113
Db	601	ATTCTCAAGCTCGTGTGATTTGCTCAGTGCAGGTCCCAATGGCATTCATCTAT	660
Qy	1114	TGATGACTGCAAGATGATTTCTTAATGAACTTTAAAGATCCTAAATCCAGTTGTATA	1173
Db	661	TGATGACTGCAAGATGATTTCTTAATGAACTTTAAAGATCCTAAATCCAGTTGTATA	720
Qy	1174	TGG 1176	
Db	721	TGG 723	

RESULT 2
BM711125
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

BM711125 679 bp mRNA linear EST 28-FEB-2002
UI-E-DX1-agv-1-12-0-UI-E1 UI-E-DX1 Homo sapiens cDNA clone
BM711125
BM711125.1 GI:19024383
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 679)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)

MEDLINE
8889548
PUBMED
COMMENT
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@iowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
Seq primer: M13 Reverse.

FEATURES

source
1. 679
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-DX1-agv-1-12-0-UI"
/tissue_type="fetal eyes"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-DX1"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-DX1 is a normalized cDNA library containing the following tissue(s): fetal eyes. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is ACAATCAGCA. This library was created for the program, Gene discovery in the Visual System, supported by National Eye Institute (NEI)."
BASE COUNT 242 a 133 c 161 g 142 t 1 others
ORIGIN

Query Match 19.4%; Score 526; DB 12; Length 679;
Best Local Similarity 99.6%; Pred. No. 4.6e-272;
Matches 676; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

1964 GAGAAATAGTACATTTTGGAAATGCACTCCGAAAGTCCAGAGAGCGCTGCTATTGG 2023
1 GAGATAGTACATTTTGGAAATGCACTCCGAAAGTCCAGAGAGCGCTGCTATTGG 60
2024 CAATTCACAGAGCGAAATGAGAGCGAAAGAGAGATGAGTGTATCATATCATC 2083
61 CAATTCACAGAGCGAAATGAGAGCGAAAGAGAGATGAGTGTATCATATCATC 120
2084 AGACAGATCAAGGCTTCTGCTAGCTAGTCTTCAACAGAGAGATTCAGGCAATTACCTC 2143
121 AGACAGATCAAGGCTTCTGCTAGCTAGTCTTCAACAGAGAGATTCAGGCAATTACCTC 180
2144 TGCATGCGGTGGAACATGGTTTCATACAACTCTTAAAGTAACTCTGGAAGTCATT 2203
181 TGCATGCGGTGGAACATGGTTTCATACAACTCTTAAAGTAACTCTGGAAGTCATT 240
2204 GACACAGACATTTGGAGAACTCTTCAATAAGATGATGAGATGGCTCTTAAGACC 2263
241 GACACAGACATTTGGAGAACTCTTCAATAAGATGATGAGATGGCTCTTAAGACC 300
2264 AAGAAATGTCATATGATGACACTAGCCAGAAAGTCTGTATCAGAACTTATGCG 2323
301 AAGAAATGTCATATGATGACACTAGCCAGAAAGTCTGTATCAGAACTTATGCG 360

QY 2324 CTCATCAACCAACCCCAATCTCAACAGATGATGATGTTCTGTGAACAAGTTGGAAAAAG 2383
DB 361 CTCATCAACCAACCCCAATCTCAACAGATGATGATGTTCTGTGAACAAGTTGGAAAAAG 420
QY 2384 GACCGAAAAACAGTCGGCAAAAGGCGAGACATACCCAGAGAAAGTAAAGAAAG 2443
DB 421 GACCGAAAAACAGTCGGCAAAAGGCGAGACATACCCAGAGAAAGTAAAGAAAG 480
QY 2444 CACTTACAAAGAAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 2503
DB 481 CACTTACAAAGAAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAG 540
QY 2504 AGAGTGTCTGAGCTGCTACCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 2563
DB 541 AGAGTGTCTGAGCTGCTACCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 600
QY 2564 ATRACTGAAAAAACAATGCAATATACATGAACTTTTTCATGATATGATGATGTT 2623
DB 601 ATRACTGAAAAAACAATGCAATATACATGAACTTTTTCATGATATGATGATGTT 660
QY 2624 ACAATGCTGGGAAATTCAG 2642
DB 661 ACAATGCTGGGAAATTCAG 679

RESULT 3
BM723807
LOCUS
DEFINITION
UI-E-B01-81x-o-04-0-UI-r1 UI-E-B01 Homo sapiens cDNA clone
BM723807
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 702)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL
MEDLINE
PUBMED
8889548

COMMENT
Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@iowa.edu
Tissue Procurement: Dr. Gregg Hageman
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).
Seq primer: M13 Reverse.
FEATURES
source
1. 702
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-B01-81x-o-04-0-UI"
/tissue_type="fetal eye"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-B01"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-B01 is a normalized cDNA library containing the following tissue(s): fetal eye. The library was constructed according to Bonaldo, Lennon and Soares,

constructed according to Bonaldo, Lennon and Soares,

Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is CCGGTACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

BASE COUNT 248 a 145 c 164 g 145 t
ORIGIN

Query Match 15.4%; Score 418; DB 12; Length 702;
Best Local Similarity 99.6%; Pred. No. 1.1e-213;

Matches 518; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1972 TAGCAGATTTTGAATGAGTCCGAGAGCGCTGCTATTGGCAATTCC 2031
DB 52 TAGCAGATTTTGAATGAGTCCGAGAGCGCTGCTATTGGCAATTCC 111
QY 2032 GAGGCGAATGAAGAGGAAAGAGATGAGTGAATCATATCATGAGACAG 2091
DB 112 GAGGCGAATGAAGAGGAAAGAGATGAGTGAATCATATCATGAGACAG 171
QY 2092 TCAGAGCCTTCTGCTAGTGTACACAGAGATTCAGCAATTAACCTGCGATGC 2151
DB 172 TCAGAGCCTTCTGCTAGTGTACACAGAGATTCAGCAATTAACCTGCGATGC 211
QY 2152 GGTGGAACATGGTTCATACAACTCTTTAAGTAACTTGAAGCTTGAACACAG 2211
DB 232 GGTGGAACATGGTTCATACAACTCTTTAAGTAACTTGAAGCTTGAACACAG 291
QY 2212 GCATTTGGAGAACTTTCATATAAGTATGATGAGATGCTTAAAGCCAAAGAAAT 2271
DB 292 GCATTTGGAGAACTTTCATATAAGTATGATGAGATGCTTAAAGCCAAAGAAAT 351
QY 2272 GTTCATATGATGACACTAGCAGAGAGTCTGTACAGAGATTCATGAGCTCATCA 2331
DB 352 GTTCATATGATGACACTAGCAGAGAGTCTGTACAGAGATTCATGAGCTCATCA 411
QY 2332 CCAACCCCAATCTCAACAGATGATGATGCTGTGAAACAAGTTGGAAGGAGCCGAA 2391
DB 412 CCAACCCCAATCTCAACAGATGATGATGCTGTGAAACAAGTTGGAAGGAGCCGAA 471
QY 2392 ACAACGTCGGCAAGGCCAGACATACCCCAAGGAAACATTAAGTGAAGCATTTACA 2451
DB 472 ACAACGTCGGCAAGGCCAGACATACCCCAAGGAAACATTAAGTGAAGCATTTACA 531
QY 2452 AGAAATTAAGAAAGTGAAGACAGAGACCCACGAATTT 2491
DB 532 AGAAATTAAGAAAGTGAAGACAGAGACCCACGAATTT 571

RESULT 4
CD216784 510 bp mRNA linear EST 20-MAY-2003
LOCUS EST Y164 Human fibroblasts senescence downregulated subtraced
DEFINITION library Homo sapiens cDNA similar to SEM3A, mRNA sequence.
ACCESSION CD216784
VERSION CD216784.1 GI:30956759
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 510)
Zhao, L., Zhang, Z., and Tong, T.
Expressed sequence tags from a human subtraced library,
downregulated in senescent fibroblasts
JOURNAL
Unpublished
COMMENT Contact: Zhao L

Department of Biochem and Molecular Biology
Peking University, Health Science Center
No.38, Xueyuan Road, Beijing 100083, P R China
Tel: 86 10 62091454
Email: to_zl@eml.com

FEATURES
source
Location/Qualifiers
1..510
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/sex="female"
/cell_type="fibroblast"
/cell_line="2BS"
/dev_stage="embryo"
/clone_lib="Human fibroblasts senescence downregulated
subtraced library"
/note="Organ: lung"

BASE COUNT 145 a 111 c 117 g 137 t
ORIGIN

Query Match 15.1%; Score 408; DB 14; Length 510;
Best Local Similarity 99.6%; Pred. No. 2.5e-208;

Matches 508; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1428 ACAATCCAGTGTTCCTATGAACAAATCCGCCAATATGATCAAAACGATTAATTATC 1487
DB 1 ACAATCCAGTGTTCCTATGAACAAATCCGCCAATATGATCAAAACGATTAATTATC 60
QY 1488 AATTATACAAATTTGCTGAGACCGAGTGAAGAGAGATGAGATGATGATGATGAT 1547
DB 61 AATTATACAAATTTGCTGAGACCGAGTGAAGAGATGAGATGATGATGATGATGAT 120
QY 1548 TTATCGGAACAGATGTGGAGCCGTTCTTAAAGTATCAATTCCTAAGAGACTTGGT 1607
DB 121 TTATCGGAACAGATGTGGAGCCGTTCTTAAAGTATCAATTCCTAAGAGACTTGGT 180
QY 1608 ATGATTTAGAGAGGTTCTGCTGAGAAAGATGACAGTTTTCGGAACCGACTGCTATT 1667
DB 181 ATGATTTAGAGAGGTTCTGCTGAGAAAGATGACAGTTTTCGGAACCGACTGCTATT 240
QY 1668 CAGCATGAGAGTTTCCACTAGCAGCAACATTAATTTGGTTCAACGGCTGGGGTTG 1727
DB 241 CAGCATGAGAGTTTCCACTAGCAGCAACATTAATTTGGTTCAACGGCTGGGGTTG 300
QY 1728 CCCAGTCCCTTTACACCGGTGTGATTTTACGGAAGCGTGTGAGTGTCTGCTG 1787
DB 301 CCCAGTCCCTTTACACCGGTGTGATTTTACGGAAGCGTGTGAGTGTCTGCTGCTG 360
QY 1788 CCGGAGCCCTTACTGTGTGGATGGTTCGATGATTTTCCACTGCA 1847
DB 361 CCGGAGCCCTTACTGTGTGGATGGTTCGATGATTTTCCACTGCA 420
QY 1848 AAGAGCGCACAAGAGACAGATTAAGAAATGAGAGCCACTGACTACGTTGAGCT 1907
DB 421 AAGAGCGCACAAGAGACAGATTAAGAAATGAGAGCCACTGACTACGTTGAGCT 480
QY 1908 TACACCATGATATACCATGCGCCACAGCC 1937
DB 481 TACACCATGATATACCATGCGCCACAGCC 510

RESULT 5
BF667677 889 bp mRNA linear EST 21-DEC-2000
LOCUS BF667677
DEFINITION 60212173P1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4278810 5',
mRNA sequence.
ACCESSION BF667677
VERSION BF667677.1 GI:11941572
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
60212173P1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4278810 5',
mRNA sequence.
JOURNAL
Unpublished
COMMENT Contact: Zhao L

Db	Accession	Version	Keywords	Source	Organism	Title	Journal	Comment
Db	225	ATTGAGAGGGCACCACGAGTGTCTGAGCTGCATTACCTCTAGAAACCTCAACAGTA						
Qy	2548	GAACCTTGCCCTAGACAACTACTGAAAAACAATGCAATATACATGAACCTTTTTCATGG						
Db	285	GAACCTTGCCCTAGACAACTACTGAAAAACAATGCAATATACATGAACCTTTTTCATGG						
Qy	2608	CATTATGTCGATGTTTACATGTCGTGGAAATTCAGCTGATGTTCCCAATTTATTAATTA						
Db	345	CATTATGTCGATGTTTACATGTCGTGGAAATTCAGCTGATGTTCCCAATTTATTAATTA						
Qy	2668	ATCCATGAGTAACTTTCTTAATAGGCT 2694						
Db	405	ATCCATGAGTAACTTTCTTAATAGGCT 431						
RESULT 7								
LOCUS	BX283116							
DEFINITION	BX283116 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4278810, mRNA sequence.							
ACCESSION	BX283116							
VERSION	BX283116.1							
KEYWORDS	EST.							
SOURCE	Homo sapiens (human)							
ORGANISM	Homo sapiens							
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
REFERENCE	1 (bases 1 to 382)							
AUTHORS	Ehbert, L., Hell, O., Hennig, S., Neubert, P., Parzsch, E., Peters, M.,							
	Radelof, U., Schneider, D. and Korn, B.							
	Human Unigeneset - RZPD3							
TITLE	Unpublished							
JOURNAL	Contact: Ina Rolfs							
COMMENT	RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH							
	Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany							
	RZPD, IMGP958A191103.							
	RZPDLIB; I.M.A.G.E. cDNA Clone Collection;							
	Human Unigeneset - RZPD3 (RZPDLIB No.972)							
	http://www.rzpd.de/cloneCards/cgi-							
	bit/showlib.pl.cgi/response?libNo=972 Contact: Ina Rolfs							
	RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH							
	Heubnerweg 6, D-14059 Berlin, Germany							
	Tel: +49 30 32639 101							
	Fax: +49 30 32639 111							
	www.rzpd.de							
	This clone is available royalty-free from RZPD.							
	contact RZPD (clone@rzpd.de) for further information. Seq primer:							
	PCMV-M13u. Primer sequence: CTTGTAAACAGCGCCAGT.							
FEATURES								
source	1..382							
	Location/Qualifiers							

BASE COUNT	106 a	85 c	94 g	97 t
ORIGIN				
Query Match	14.1%	Score 382	DB 13	Length 382

	Best Local Similarity	100.0%;	Pred. No. 2.6e-194;	
	Matches	382;	Conservative 0;	Mismatches 0;
			Indels 0;	Gaps 0;
QY	1595	AAGAGACTTGGTATGATTTAGAAAGAGTTCTCTGTGAAAGAAATGACAGTTTTCGGGAA	1654	
Db	1	AAGAGACTTGGTATGATTTAGAAAGAGTTCTCTGTGAAAGAAATGACAGTTTTCGGGAA	60	
QY	1655	CCGACTCTATTTCAGCAATGAGAGCTTCCACTAGACAGCAACAATAATATTTGGTTCA	1714	
Db	61	CCGACTCTATTTCAGCAATGAGAGCTTCCACTAGACAGCAACAATAATATTTGGTTCA	120	
QY	1715	ACGGCTGGGGTTGCCACAGCTCCCTTTACACCGGTGATTTTAAACGGGAAAGCGTGTCT	1774	
Db	121	ACGGCTGGGGTTGCCACAGCTCCCTTTACACCGGTGATTTTAAACGGGAAAGCGTGTCT	180	
QY	1775	GAGGTGGCTCGCCGACAGACCCCTTACTGTCTTGGGATGGTCTGTGATTTTCGGTAT	1834	
Db	181	GAGGTGGCTCGCCGACAGACCCCTTACTGTCTTGGGATGGTCTGTGATTTTCGGTAT	240	
QY	1835	TTTCCACTCGAAGAGACGACACAGACGCAACAAGATATTAAGAAATGGAGACCCACTGACT	1894	
Db	241	TTTCCACTCGAAGAGACGACACAGACGCAACAAGATATTAAGAAATGGAGACCCACTGACT	300	
QY	1895	CACGTTCAGACTTACACCATGATTAATCACCATGGCCACAGCCCTGAAGAGAAATCATC	1954	
Db	301	CACGTTCAGACTTACACCATGATTAATCACCATGGCCACAGCCCTGAAGAGAAATCATC	360	
QY	1955	TATGTGTAGAGAAATGATGACA	1976	
Db	361	TATGTGTAGAGAAATGATGACA	382	

RESULT 8	
HSMB084323/c	
ID	HSMB084323
standard:	RNA; EST; 721 BP.
XX	
AC	BX510254;
XX	
SV	BX510254.1
XX	
DT	09-MAY-2003 (Rel. 75, Created)
DT	09-MAY-2003 (Rel. 75, Last updated, Version 1)
XX	
DE	Homo sapiens mRNA; EST DKFZp686L1998_s1 (from clone DKFZp686L1998)
XX	
KW	EST; expressed sequence tag.
XX	
OS	Homo sapiens (human)
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia
OC	Eutheria; Primates; Catarrhini; Homidae; Homo.
XX	
LN	[1]
RP	1-721
RA	Bahr A., Lauber J., Mewes H.W., Weil B., Amid C., Oanger A., Fodo G.,
RA	Han W., Wiemann S.;
RT	Submitted (09-MAY-2003) to the EMBL/GenBank/DBJ databases.
RL	MRS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY
CC	
CC	This is the 3' sequence of the clone insert
CC	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
CC	Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
CC	sequenced by Oligen (Hilden/Germany) within the cDNA sequencing
CC	consortium of the German Genome Project.
CC	rt1 sequence also available.
CC	This clone (DKFZp686L1998) is available at the RZPD in Berlin.
CC	Please contact the RZPD: Ressourcenzentrum, Heubenerweg 6,
CC	14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de
XX	
EH	Key
EH	Location/Qualifiers
EH	
EH	1. 721
EH	source
EH	/db xref="taxon:9606"
EH	

```

FT      /mol_type="mRNA"
FT      /organism="Homo sapiens"
FT      /clone_id="DKFZp686L198"
FT      /clone_1ib="686 (synonym: hlcc3). Vector pSport1_Sfi; host
FT      DH10B; sites SfiI + SfiIb"
FT      /dev_stage="adult"
FT      /tissue_type="CDNA-collection"
XX      SQ      Sequence 721 BP; 192 A; 141 C; 121 G; 267 T; 0 other;

Query Match      12.0%; Score 324; DB 2; Length 721;
Best Local Similarity 100.0%; Pred. No. 6.8e-163;
Matches 324; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2377 GAAAAGGAGCCGAAACCAAGTCGCGAAGGCGAGACATACCCGAGAACATACAA 2436
DB      690 GAAAAGGAGCCGAAACCAAGTCGCGAAGGCGAGACATACCCGAGAACATACAA 631
QY      2437 ATGGAAGCATTACAGAAATTAAGAAAGTGAAGACAGAGAGCCGAAATTTGAGAG 2496
DB      630 ATGGAAGCATTACAGAAATTAAGAAAGTGAAGACAGAGAGCCGAAATTTGAGAG 571
QY      2497 GGCACCCGAGAGTGTGAGCTGAGCTTACCTCTAGAAAGCTCAACAAAGTAACTTGC 2556
DB      570 GGCACCCGAGAGTGTGAGCTGAGCTTACCTCTAGAAAGCTCAACAAAGTAACTTGC 511
QY      2557 CTAGCAATTAATCTGAAACCAATGCAATATATATATATATATATATATATATAT 2616
DB      510 CTAGCAATTAATCTGAAACCAATGCAATATATATATATATATATATATATATATAT 451
QY      2617 GATGTTTACATATGTCGGAATTCAGCTGAGTTCACCAATTAATTAATTAATTCAT 2676
DB      450 GATGTTTACATATGTCGGAATTCAGCTGAGTTCACCAATTAATTAATTAATTCATGAG 391
QY      2677 TAACCTTCTAATAGCTTTT 2700
DB      390 TAACCTTCTAATAGCTTTT 367

RESULT 9
BM670627/c
LOCUS      BM670627      670 bp      mRNA      linear      EST 27-FEB-2002
DEFINITION      UI-E-DX1-egv-1-12-0-UI.s1 UI-E-DX1 Homo sapiens cDNA clone.
ACCESSION      BM670627
VERSION      BM670627.1 GI:18980524
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
REFERENCE      Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Bonaldo, M.F., Lennon, G. and Soares, M.B.
TITLE      Normalization and subtraction: two approaches to facilitate gene
JOURNAL      Tissue Procurement: Dr. Gregg Hageman
MEDLINE      cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
PUBMED      DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
COMMENT      Clone Distribution: Researchers may obtain clones from Research
sequence: 1-25, >AT_rich#low_complexity (matched complement)

```

```

FEATURES
source
Seq primer: M13 Forward
POLYA=yes.
Location/Qualifiers
1. 670
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-DX1-egv-1-12-0-UI"
/tissue_type="fetal eyes"
/dev_stage="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_1ib="UI-E-DX1"
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site 1: EcoR I; Site 2: Not I;
UI-E-DX1 is a normalized cDNA library containing the
following tissue(s): fetal eyes. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
AGAAATCAAGA. This library was created for the program, Gene
Discovery in the Visual System, supported by National Eye
Institute (NEI).
TAG_LIB=UI-E-DX1
TAG_TISSUE=human fetal eyes
TAG_SEQ=AGAAATCAAGA"
BASE COUNT      181 a      128 c      115 g      246 t
ORIGIN

Query Match      11.6%; Score 315; DB 12; Length 670;
Best Local Similarity 100.0%; Pred. No. 4.9e-158;
Matches 315; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2395 ACGTCGGCAAGGCGAGACATACCCGAGGAGAGTAAAGAGCACTTACAGA 2454
DB      670 ACGTCGGCAAGGCGAGACATACCCGAGGAGAGTAAAGAGCACTTACAGA 611
QY      2455 AAATAGAAAGGTGAAGAACAGAGAGACCCAGAAATTTGAGAGGACCCAGAGTGTG 2514
DB      610 AAATAGAAAGGTGAAGAACAGAGAGACCCAGAAATTTGAGAGGACCCAGAGTGTG 551
QY      2515 AGCTGCAATTAATCTGAAACCTCAACAGTGAAGAACTTCTAGACATTAATCTGAAA 2574
DB      550 AGCTGCAATTAATCTGAAACCTCAACAGTGAAGAACTTCTAGACATTAATCTGAAA 491
QY      2575 AACAAAGCAATTAATCAATTTTTCATGAGGATTAATGAGTAAAGTGAAGG 2634
DB      490 AACAAAGCAATTAATCAATTTTTCATGAGGATTAATGAGTAAAGTGAAGG 431
QY      2635 AAATTCAGCTGAGTTCACCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2694
DB      430 AAATTCAGCTGAGTTCACCAATTAATTAATTAATTAATTAATTAATTAATTAATTA 371
QY      2695 TTTTCTCTAATACC 2709
DB      370 TTTTCTCTAATACC 356

RESULT 10
BF667143
LOCUS      BF667143      839 bp      mRNA      linear      EST 21-DEC-2000
DEFINITION      602121928F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4278967 5',
ACCESSION      BF667143
VERSION      BF667143.1 GI:11941038
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens

```

REFERENCE 1 (bases 1 to 839)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: ATCC

CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Plate: LNCM103 row: h column: 08
 High quality sequence stop: 633.

FEATURES

source

1. .839

Location/Qualifiers

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:427867"
 /tissue_type="primitive neuroectoderm"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: brain; Vector: pDNR-LIB (Clontech); Site_1: SfiI (ggcgctggcc); Site_2: SfiI (ggccattggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CAGCGCCATTAGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGCGCCGACATG-dT(30)-BH-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT

299 a 129 c 195 g 216 t

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 2,4e-142;

Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2424 GGAACGTATCAATGGAACCTTCAAGAAATAGAAAGTAGAAACAGAGAGACC 2483
 |||||
 DB 73 GGAACGTATCAATGGAACCTTCAAGAAATAGAAAGTAGAAACAGAGAGACC 132
 |||||
 QY 2484 ACGAATTTGAGAGGACCCAGAGAGTCTGAGCTGATTACCTTAGAAACCTCAACA 2543
 |||||
 DB 133 ACGAATTTGAGAGGACCCAGAGAGTCTGAGCTGATTACCTTAGAAACCTCAACA 192
 |||||
 QY 2544 AGTAGAACTTGGCTTACCAATATCTGAAAAAACAATGCAATATACATGAACCTTTTTC 2603
 |||||
 DB 193 AGTAGAACTTGGCTTACCAATATCTGAAAAAACAATGCAATATACATGAACCTTTTTC 252
 |||||
 QY 2604 ATGGCAATTAATGAGATTTTACCAATGTTGGAAATCAGTGAAGTCCACCAATTTTAA 2663
 |||||
 DB 253 ATGGCAATTAATGAGATTTTACCAATGTTGGAAATCAGTGAAGTCCACCAATTTTAA 312
 |||||
 QY 2664 TTAATTCATGATTAATCTTCTTAATAGGCTTTTCTTAAATACC 2709
 |||||
 DB 313 TTAATTCATGATTAATCTTCTTAATAGGCTTTTCTTAAATACC 358
 |||||

RESULT 11
 BQ018643/c 747 bp mRNA linear EST 27-MAR-2002
 LOCUS UI-H-DH1-awu-m-14-0-UI-81 NCI_CGAP_DHI Homo sapiens cDNA clone
 DEFINITION IMAGE:5823925 3', mRNA sequence.
 ACCESSION BQ018643
 VERSION BQ018643.1 GI:19753920
 KEYWORDS EST.

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 747)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: ATCC

CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 The following repetitive elements were found in this cDNA sequence: 1-25, >AT rich#low_complexity (matched complement)
 Seg primer: M13 FORWARD
 POLYA=Yes.

FEATURES

source

1. .747

Location/Qualifiers

/organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5823925"
 /tissue_type="Metastatic Chondrosarcoma"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="NCI CGAP DHI"
 /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP_DHI is a normalized cDNA library containing the following tissue(s): VS-8 Cell line from Metastatic Chondrosarcoma in Lung. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AAGTATTCG.

BASE COUNT

193 a 143 c 134 g 275 t

ORIGIN

Query Match

Best Local Similarity 99.7%; Pred. No. 2.9e-141;

Matches 334; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2375 TGGAAAAGGAGCCGAAAAACAAGTGGGCAAGCCAGCATATCCAGGAAACAGTAAC 2434
 |||||
 DB 690 TGGAAAAGGAGCCGAAAAACAAGTGGGCAAGCCAGCATATCCAGGAAACAGTAAC 631
 |||||
 QY 2435 AAATGGAACCTTACAGAAAAATAGAAAGTAGAAACAGAGAGCCAGCAATTTGAG 2494
 |||||
 DB 630 AAATGGAACCTTACAGAAAAATAGAAAGTAGAAACAGAGAGCCAGCAATTTGAG 571
 |||||
 QY 2495 AGGCAACCCAGAGAGTCTGAGCTGATTACCTCTTGAAGAACTCAACAAAGTAGAACTT 2554
 |||||
 DB 570 AGGCAACCCAGAGAGTCTGAGCTGATTACCTCTTGAAGAACTCAACAAAGTAGAACTT 511
 |||||
 QY 2555 GCTTAGACATATACGAGAAAAACAATGCAATATACATGAACCTTTTTCATGAGCATTTATG 2614
 |||||
 DB 510 GCTTAGACATATACGAGAAAAACAATGCAATATACATGAACCTTTTTCATGAGCATTTATG 451
 |||||
 QY 2615 TGGATGTTTACATGATGTTGGAATTCAGTGAAGTCCACCAATTAATTAATCATG 2674
 |||||

Db 450 TGGATGTTTACATGGTGGAAATTCAGCTGACCTCACCATTATTAATTAATCATTG 391

QY 2675 AGTACTTCTCTAATAGGCTTTTCTTCTAATACC 2709

Db 390 AGTAACTTCTCTAATAGGCTTTTCTTCTAATACC 356

RESULT 12

BF700780 884 bp mRNA linear EST 22-DEC-2000

LOCUS 602128585F1 NIH_MGC_56 Homo sapiens cDNA clone IMAGE:4285532 5'

DEFINITION mRNA sequence.

ACCESSION BF700780

VERSION BF700780.1 GI:11986188

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabds-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: L1CM1120 row: 1 column: 21
High quality sequence stop: 676.

FEATURES

source

1..884

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4285532"

/tissue_type="primitive neuroectoderm"

/lab_host="DH10B (T1 phage-resistant)"

/clone_11b="NIH_MGC_56"

/note="Organ: brain; Vector: pDNR-LIB (Clontech); site_1: SfiI (ggcgccatcgcc); Site 2: SfiI (ggcgatcgcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATCTAGAGCGCGAGCGCGACATG-3' and 3' adaptor sequence: 5'-ATCTAGAGCGCGAGCGCGACATG-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

BASE COUNT 280 a 185 c 225 g 194 t

ORIGIN

Query Match 10.1%; Score 274; DB 10; Length 884;

Best Local Similarity 100.0%; Pred. No. 7.6e-136; Indels 0; Gaps 0;

Matches 274; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1916 GATAATCACCATGCGACAGCCCTGAAGAGAGATCATCTATGTTAGATAGTATG 1975

Db 52 GATATACACATGCGACAGCCCTGAAGAGAGATCATCTATGTTAGATAGTATG 111

QY 1976 ACATTTTGAATGAGTCCGAAAGTGCAGAGAGCGCTGCTTATTTGGCAATTCAGAG 2035

Db 112 ACATTTTGAATGAGTCCGAAAGTGCAGAGAGCGCTGCTTATTTGGCAATTCAGAG 171

QY 2036 CGAATGAGAGGAGAAAGAGATCAGAGTGAATCATCTATCTCAGAGAGATGCA 2095

Db 172 CGAATGAGAGGAGAAAGAGATCAGAGTGAATCATCTATCTCAGAGAGATGCA 231

QY 2096 GGCCTTCTGCTACGATGCTTACACAGAGATTCAGGCATTACTTGCATGCGGTG 2155

Db 232 GGCCTTCTGCTACGATGCTTACACAGAGATTCAGGCATTACTTGCATGCGGTG 291

QY 2156 GAACATGGGCTCATCAAACTCTCTTAAAGTAA 2189

Db 292 GAACATGGGCTCATCAAACTCTCTTAAAGTAA 325

RESULT 13

HS08084322

ID HS08084322 standard; RNA; EST; 535 BP.

AC BX491494;

XX BX491494.1

SV BX491494.1

DT 09-MAY-2003 (rel. 75, Created)

DT 09-MAY-2003 (rel. 75, Last updated, Version 1)

DE Homo sapiens mRNA; EST DKFZp686L1998_r1 (from clone DKFZp686L1998)

XX EST; expressed sequence tag.

XX Homo sapiens (human)

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

XX [1]

RP 1-535

RA Bahr A., Lauber J., Mewes H.W., Weil B., Amid C., Osanger A., Fobo G., Han M., Wiemann S.;

RT Submitted (07-MAY-2003) to the EMBL/GenBank/DBJ databases.

RL MIPB, Ingolstaedter Landstr.1, D-85764 Neuberg, GERMANY

XX This is the 5' sequence of the clone insert

CC Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing consortium of the German Genome Project.

CC S1 sequence also available.

CC This clone (DKFZp686L1998) is available at the RZPD in Berlin.

CC Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de

XX

XX Key Location/Qualifiers

XX source

1..535

/db_xref="taxon:9606"

/mol_type="mRNA"

/organism="Homo sapiens"

/clone="DKFZp686L1998"

/clone_11b="686 (synonym: h1c3). Vector pSport1_Sfi; host DH10B; sites SfiI + SfiIb"

/dev_stage="adult"

/tissue_type="cDNA-collection"

Sequence 535 BP; 135 A; 108 C; 119 G; 173 T; 0 other;

Query Match 9.8%; Score 265; DB 2; Length 535;

Best Local Similarity 100.0%; Pred. No. 4.8e-111; Indels 0; Gaps 0;

Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 TTTCAGATGTTGAACTTCTCTGCGGACACATACAGAGAAAGACTAAAGCAGCA 177

Db 271 TTTCAGATGTTGAACTTCTCTGCGGACACATACAGAGAAAGACTAAAGCAGCA 330

QY 178 AGGACCTACAGCGCTGCTCAGATGAGTGGTGAATCTGATGTTGCTCTTTCTGAGG 237

Db 331 AGGACCTACAGCGCTGCTCAGATGAGTGGTGAATCTGATGTTGCTCTTTCTGAGG 390

QY 238 AGTATTACTTACAGCAAGCAAACTATCAGATGGAGAAACAATGTGCCAAGCTGAA 297

Accession	Sequence	Position
Db	391 AGTATTAATTACGACAAGACCAATATACAGATGTGGAAAGAACATGTGTCCCAAGCTGAA	450
Oy	298 ATTATCTTACCAAGAAATGTTGGATCCACATGTGATCACTTTCOANTGCTTGCCCA	357
Db	451 ATTATCTTACCAAGAAATGTTGGATCCACATGTGATCACTTTCOANTGCTTGCCCA	510
Oy	358 CAGCTCCAGTTATCATCTACCTTCCTT	382
Db	511 CAGCTCCAGTTATCATCTACCTTCCTT	535

RESULT 14	LOCUS	DEFINITION	LOCUS	DEFINITION
BQ447821/c	BQ447821	692 bp	mRNA	linear
	UI-H-EU1-bag-c-21-0-UI-s1	NCI CGAP C1	Homo sapiens	CDNA clone
	UI-H-EU1-bag-c-21-0-UI 3'			mRNA sequence.

ACCESSION	BQ447821	
VERSION	BQ447821.1	GI:212509333
KEYWORDS	EST.	
SOURCE	Homo sapiens (human)	
ORGANISM	Homo sapiens	

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Unpublished
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 1 (bases 1 to 692)
 Mammalia; Eutheria; Primates; Carnivora; Hominiidae; Homo.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained
from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
PolyA=yes.

FEATURES	Location/Qualifiers
source	1. .692

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-EU1-Bag-c-21-0-UI"
/tissue_type="Osteoarthritic Cartilage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_1ib="NCI_CGAP_Ct1"
/name="Organ: Knee; Vector: pRTT3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
NCI_CGAP_Ct1 is a normalized cDNA library containing the
following tissue(s): Osteoarthritic Cartilage The library
was constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pRTT3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
TGATCAGCGT.
TAG_lib=UI-H-EU1
TAG_tissue=osteoarthritic cartilage
TAG_SEQ=TGATCAGCGT"
181 a 136 c 119 g 255 t 1 others

```

Query March 260; DB 13; Length 692; Best Local Similarity 100.0%; Pred. No. 2.6e-126; Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	2450	AAAAAAATAGAAAGGTAGAAAACAGAGAGCCACAAATTTGAGAGGACCCAGAGT	2509
Db	608	CAAGAAATTAAGAAAGGTAGAAACAGAGAGCCACAAATTTGAGAGGACCCAGAGT	549
Qy	2510	GTCTGAGCTGCATTACCTCTAGAAACCTCAAACAGTAGAAACTTGCTTGACAAATACT	2569
Db	548	GTCTGAGCTGCATTACCTCTAGAAACCTCAAACAGTAGAAACTTGCTTGACAAATACT	489
Qy	2570	GGAAAAACAATGCAATATACATGAACCTTTTTCATGGCAATTATGAGATGTTACATG	2629
Db	488	GGAAAAACAATGCAATATACATGAACCTTTTTCATGGCAATTATGAGATGTTACATG	429
Qy	2630	GTGGGAAATTCAGCTGAGTTCCACCAATTTAAATTAATTCATGAGTAATCTTCTAAT	2689
Db	428	GTGGGAAATTCAGCTGAGTTCCACCAATTTAAATTAATTCATGAGTAATCTTCTAAT	369
Qy	2690	AGGCTTTTTCCTAATACC	2709
Db	368	AGGCTTTTTCCTAATACC	349

[illegible]

REFERENCE
AUTHORS

1 (pages 1 to 355)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Biontes, M.R., Negai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, R.F., Goldmann, G.R., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE	JOURNAL	PUBMED	COMMENT
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)	20202663	
		10737800	Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
<http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL2&t2=IL2-B70734-1>
 0A1000-178-C03&t3=2000-10-04&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 355.
 Location/Qualifiers

```

source
1. .355
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_1b="BT0734"
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Ludwig Institute for Cancer Research) profiles into the puc18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
BASE COUNT      103 a      83 c      64 g      105 t

```

ORIGIN

Query Match 8.0%; Score 218; DB 10; Length 355;
 Best Local Similarity 100.0%; Pred. No. 1.1e-105; Indels 0; Gaps 0;
 Matches 218; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 649 TGAGGACATATTTTAACTGAGAACTACATTTTGAAAACGGCGTGGAGAGTCC 708
 Db 355 TGAGGACATATTTTAACTGAGAACTACATTTTGAAAACGGCGTGGAGAGTCC 296
 QY 709 AATATGACCTTAAGCTCTGACAGCATCCCTTTAATATGATGAGATTTACTCTGGAAC 768
 Db 295 AATATGACCTTAAGCTCTGACAGCATCCCTTTAATATGATGAGATTTACTCTGGAAC 236
 QY 769 TCGAGCTGATTTATGAGGCGAGACTTCTCTATCTTCCGAACTCTGGGACACACACC 828
 Db 235 TCGAGCTGATTTATGAGGCGAGACTTCTCTATCTTCCGAACTCTGGGACACACACC 176
 QY 829 AATCAGACAGACAGCATGATTTCCAGTGGCTCAATG 866
 Db 175 AATCAGACAGACAGCATGATTTCCAGTGGCTCAATG 138

RESULT 16 231 bp mRNA linear EST 06-SEP-1995
 LOCUS EST6074 Human lung Homo sapiens cDNA 5' end similar to semaphorin
 DEFINITION III (HT:3407), mRNA sequence.

ACCESSION T29595
 VERSION T29595.1 GI:611693
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AUTHORS Adams,M.D., Kerlavage,A.R., Fleischmann,R.D., Fuldner,R.A., Bult
 C.J., Lee,N., Kirkness,E.F., Weinstock,K.G., Gocayne,J.D., White
 O., Sutton,G., Blake,J.A., Brandon,R.C., Chiu,M.-W., Clayton,R.A.,
 Cline,R.T., Cotton,M.D., Batle-Hughes,J., Fine,L.D., Fitzgerald,
 L.M., Fitzhugh,W.M., Fritchman,J.L., Geoghegan,N.S.M., Glodok,A.,
 Gnehm,C.L., Hanna,M.C., Hedblom,E., Hinklejr,P.S., Kelley,J.M.,
 Klinek,K.M., Kelley,J.C., Liu,L.-I., Marmaros,S.M., Merrick,J.M.,
 Moreno-Palanges,R.F., McDonald,L.A., Nguyen,D.T., Pellegrino,S.M.,
 Phillips,C.A., Ryder,S.E., Scott,J.L., Soudk,D.M., Shirley,R.,
 Small,K.V., Spriggs,T.A., Ulteback,T.R., Weidman,J.F., Li,Y.,
 Bednarek,D.P., Cao,L., Cepeda,W.A., Coleman,T.A., Collins,E.-J.,
 Dimde,D., Feng,P., Fertle,A., Fischer,C., Hastings,G.A., He,W.-W.,
 Hu,J.-S., Greene,J.M., Gruber,J., Hudson,P., Kim,A., Kozak,D.L.,
 Kunesh,C., Ji,H., Li,H., Meisner,P.S., Olsen,H., Raymond,L., Wei
 Y.-F., Wang,T., Xu,C., Yu,G.-L., Ruben,S.M., Dillon,P.J., Fannon
 M.R., Rosen,C.A., Hesselink,W.A., Fields,C., Frazer,C.M. and
 Venter,J.C.

TITLE Initial Assessment of Human Gene Diversity and Expression Patterns
 JOURNAL Nature 377, 3-174 (1995)
 MEDLINE 96036280
 PUBMED 7566098
 COMMENT Contact: Venter, JC
 The Institute for Genomic Research
 932 Clopper Rd, Gaithersburg, MD 20878
 Tel: 3018699056
 Fax: 3018699423

FEATURES
 source
 Email: tdbinfo@db.tigr.org
 For clone availability, this EST, please contact the TIGR Database
 (tdbinfo@db.tigr.org)
 Seq primer: M3 Reverse.

Location/Qualifiers
 1..231
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="ATCC (inhost):106785"

/db_xref="taxon:9606"
 /clone_11b="Human Lung"
 /note="Organ: Lung"
 BASE COUNT 65 a 48 c 56 g 62 t
 ORIGIN

Query Match 7.9%; Score 213; DB 14; Length 231;
 Best Local Similarity 100.0%; Pred. No. 4.9e-103; Indels 0; Gaps 0;
 Matches 213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1123 GCAGATGATTTCTTAATGAACCTTAAGATCTTAATAATCCAGTTGATATGAGTGT 1182
 Db 1 GCAGATGATTTCTTAATGAACCTTAAGATCTTAATAATCCAGTTGATATGAGTGT 60
 QY 1183 TACGACTTCAGTAACTTTTCAAGGGATCGCGTGATATGATATGAGTGT 1242
 Db 61 TACGACTTCAGTAACTTTTCAAGGGATCGCGTGATATGATATGAGTGT 120
 QY 1243 GAGAAAGGTGTTCTTGTCCATATGCCCACAGGATGACCACTATCATATGAGTGTCC 1302
 Db 121 GAGAAAGGTGTTCTTGTCCATATGCCCACAGGATGACCACTATCATATGAGTGTCC 180
 QY 1303 TTATCAGGAAGAGTCCCTTATCCAGGCGCAGG 1335
 Db 181 TTATCAGGAAGAGTCCCTTATCCAGGCGCAGG 213

RESULT 17 328 bp mRNA linear EST 17-DEC-1995
 LOCUS A1671991/c
 DEFINITION wb72c04.x1 NCI CGAP GC6 Homo sapiens cDNA clone IMAGE:2311206 3'
 similar to TR:Q14563 Q14563 SEMAPHORIN-III., mRNA sequence.

ACCESSION A1671991
 VERSION A1671991.1 GI:4851722
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 JOURNAL Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgsapb@mail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
 R. Emmert-Buck, M.D., Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNI, at:
 www.bio.lnlnl.gov/bbrp/image/image.html
 Insert length: 358 Std Error: 0.00
 Seq primer: -40bp from Gibco.

FEATURES

source
 Location/Qualifiers
 1..328
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:2311206"
 /tissue_type="pooled germ cell tumors"
 /lab_host="DH10B"
 /clone_11b="NCI CGAP GC6"
 /note="Vector: pRT3D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA
 from the normalized library NCI-CGAP GC4 was prepared, and
 ss circles were made in vitro. Following HAP purification,
 this DNA was used as tracer in a subtracting hybridization
 reaction. The driver was PCR-amplified cDNAs from a pool
 of 5,000 clones made from the same library (clonoids)

BASE COUNT 1257096-1258631, 1469064-1470983, and 1475592-1476743).
 Subtraction by Bento Soares and M. Fatima Bonaldo. "

Query Match 7.5%; Score 203; DB 9; Length 328;
 Best Local Similarity 100.0%; Pred. No. 1,4e-97;
 Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 807 GAACTCTGGGACACACCAATGAGACAGACAGCATGATTCAGGTCATG 866
 DB 328 GAACTCTGGGACACACCAATGAGACAGACAGCATGATTCAGGTCATG 269
 Y 867 ATCAAGATTCATGATGACCACTCATCTCAGAGATGCAATCTGAAGATGACAAG 926
 DB 268 ATCAAGATTCATGATGACCACTCATCTCAGAGATGCAATCTGAAGATGACAAG 209
 Y 927 TATACCTTTCTTCCTCGGAAATGCAATAGAGAACTCTGAAAAGCTACTACG 986
 DB 208 TATACCTTTCTTCCTCGGAAATGCAATAGAGAACTCTGAAAAGCTACTACG 149
 Y 987 CTGAATAGGTCAGATGCAAG 1009
 DB 148 CTGAATAGGTCAGATGCAAG 126

RESULT 18 210 bp mRNA linear EST 17-APR-1995
 LOCUS R19784
 DEFINITION Y928612.r1 Soares infant brain INIB Homo sapiens cDNA clone
 IMAGE:33664.5; similar to SP:A49069 A49069 COLLA5IN - ; mRNA
 sequence.

ACCESSION R19784
 VERSION R19784.1 GI:774418
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 210)
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucada, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, B., Rohlfing, T., Soares, M., Tan, F., Trevisan, E., Waterston, R., Williamson, A., Wohlmann, P., and Wilson, R.
 The Washu-Merck EST Project
 Unpublished

TITLE Unpublished
 JOURNAL
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu

Insert Size: 1370
 High quality sequence stops: 127 Source: IMAGE Consortium, LNL
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 1370 Std Error: 0.00
 Seg primer: M13Rpl
 High quality sequence stop: 127.

FEATURES
 source Location/Qualifiers

1..210
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:406011"
 /db_xref="taxon:9606"
 /clone="IMAGE:33664"
 /sex="female"
 /dev_stage="73 days post natal"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_lib="Soares infant brain INIB"
 /note="Organ: whole brain; Vector: Lactid BA; Site: 1: Not
 1; Site: 2: Hind III; 1st strand cDNA was primed with a Not
 I - oligo (dT) primer [5'

BASE COUNT 65 a 44 c 53 g 46 t
 Query Match 7.2%; Score 196; DB 14; Length 210;
 Best Local Similarity 100.0%; Pred. No. 7.3e-94;
 Matches 196; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 2000 TCGCAGAGCCCTGCTTATGCAATTCAGAGCGCAATGAAGGAAAGAG 2059
 DB 1 TCGCAGAGCCCTGCTTATGCAATTCAGAGCGCAATGAAGGAAAGAG 60
 Y 2060 ATCAGAGTGCATGATCATTCATCAGACATCAAGGCTTCTGCTACTGCTACCA 2119
 DB 61 ATCAGAGTGCATGATCATTCATCAGACATCAAGGCTTCTGCTACTGCTACCA 120
 Y 2120 CAGAGGATTCAGGCAATTCCTGCAATGCGGTGGAACATGGTTCATCAACTCTT 2179
 DB 121 CAGAGGATTCAGGCAATTCCTGCAATGCGGTGGAACATGGTTCATCAACTCTT 180
 Y 2180 CTTAAGGTAAACCTGG 2195
 DB 181 CTTAAGGTAAACCTGG 196

RESULT 19 914 bp mRNA linear EST 23-MAY-2003
 LOCUS AL529332
 DEFINITION AL529332 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
 CDNA clone CSODD002YV16 5-PRIME, mRNA sequence.

ACCESSION AL529332
 VERSION AL529332.2 GI:31067175
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 914)
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 Full-length cDNA libraries and normalization

TITLE Unpublished
 JOURNAL
 COMMENT On Feb 13, 2001 this sequence version replaced gi:12792825.

Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by life technologies, a division of
 invitrogen. This sequence belongs to sequence cluster 4349.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CSODD002D0808P1&cluster=4349.r. Contact :
 Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ invitrogen Corporation 1600
 Paraday Avenue Genoscope sequence ID : CSODD002D0808P1.

FEATURES
 source Location/Qualifiers

1..914
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSODD002YV16"
 /cissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo (dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and BclI V
 sites of the pCMVSPORT 6 vector. Library was normalized."
 BASE COUNT 279 a 198 c 183 g 253 t
 ORIGIN

Query Match 7.2%; Score 194; DB 9; Length 914;
 Best Local Similarity 100.0%; Pred. No. 1.3e-92;
 Matches 194; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 TTTCAGATTGTTGAACCTTCTGCGCCGCAATACAGAGAGAAAGACTAAAGCAGCA 177
 DB 479 TTTCAGATTGTTGAACCTTCTGCGCCGCAATACAGAGAGAAAGACTAAAGCAGCA 538
 QY 178 AGGACCTACAGGCTGCGAGCATGGGCTGTTACATGATGCTGCTTTCTGGGG 237
 DB 539 AGGACCTACAGGCTGCGAGCATGGGCTGTTACATGATGCTGCTTTCTGGGG 598
 QY 238 AGTATTACTTACAGCAAGCAAACTATCAGATGGGAAAGCAATGTCCTCAAGCTGAA 297
 DB 599 AGTATTACTTACAGCAAGCAAACTATCAGATGGGAAAGCAATGTCCTCAAGCTGAA 658
 QY 298 ATTATCTCTCAAG 311
 DB 659 ATTATCTCTCAAG 672

RESULT 20

BE552070/c 325 bp mRNA linear EST 10-AUG-2000
 LOCUS hy03d11.x1 NCI CGAP GC6 Homo sapiens CDNA clone IMAGE:3196245 3'
 DEFINITION similar to TR:Q14563 Q14563 SEMAPHORIN-III.; mRNA sequence.

ACCESSION BE552070
 VERSION BE552070.1 GI:9793762
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncigap.
 Tumor Gene Index
 Unpublished
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.

CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
 Seg primer: -40up from Gibco.
 Location/Qualifiers

FEATURES

1..325
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3196245"
 /issue_type="pooled germ cell tumors"
 /lab_host="DH10B"
 /clone_lib="NCI-CCGAP GC6"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI-CCGAP GC4 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clone IDs 1257096-1258631, 1469064-1470983, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo."
 BASE COUNT 81 a 63 c 67 g 114 t

Query Match 5.0%; Score 135; DB 10; Length 325;

Best Local Similarity 100.0%; Pred. No. 7.6e-61;
 Matches 135; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 807 GAACCTTGGGACACACCAACCAATCAGACAGACACACATTCAGAGGCTCATG 866
 DB 325 GAACCTTGGGACACACCAACCAATCAGACAGACACACATTCAGAGGCTCATG 266
 QY 867 ATCCAAAGTTCATTAGTGGCCCACTCATCTGAGAGTGACCAATCTGAAGTGA 926
 DB 265 ATCCAAAGTTCATTAGTGGCCCACTCATCTGAGAGTGACCAATCTGAAGTGA 206
 QY 927 TATACCTTTTCTTCC 941
 DB 205 TATACCTTTTCTTCC 191

RESULT 21

AA424316 430 bp mRNA linear EST 16-OCT-1997
 LOCUS zv90d08.r1 Soares NHMPu_S1 Homo sapiens CDNA clone IMAGE:767055
 DEFINITION 5', mRNA sequence.

ACCESSION AA424316
 VERSION AA424316.1 GI:2103286
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 430)
 Hillier, L., Allen, M., Bowles, J., Dubuque, T., Geisel, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Merra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wyllie, T., Waterston, R. and Wilson, R.
 WashU-Merck EST Project 1997

Unpublished
 CONTACT: Wilson RK
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810

Email: est@watson.wustl.edu
 This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seg primer: -28m3 rev2 ET from Amersham
 High quality sequence stop: 413.
 Location/Qualifiers

FEATURES

1..430
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:767055"
 /issue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
 /lab_host="DH10B"
 /clone_lib="Soares NHMPu_S1"
 /note="Organ: mixed (see below); Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NbH, pregnant uterus 2NbH, and fetal heart 2NbH19w) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."
 BASE COUNT 152 a 82 c 78 g 118 t

Query Match 4.4%; Score 120; DB 9; Length 430;
 Best Local Similarity 100.0%; Pred. No. 1e-52;
 Matches 120; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2485 CGAATTGAGAGGACCCAGAGTGTCTAGCTGACTTACTCTAGAACCTTCAACAA 2544
 Db 37 CGAATTGAGAGGACCCAGAGTGTCTAGCTGACTTACTCTAGAACCTTCAACAA 96
 QY 2545 GTGAAACTTGCTTAGACATTAAGTGAAGAAACAAATGCAATATACCTGACTTTTTC 2604
 Db 97 GTGAAACTTGCTTAGACATTAAGTGAAGAAACAAATGCAATATACCTGACTTTTTC 156

RESULT 22
 BUI86364 896 bp mRNA linear EST 04-SEP-2002
 LOCUS AGENCOURT_7953921 NIH_MGC_68 Homo sapiens cDNA clone IMAGE:6010676
 DEFINITION 5', mRNA sequence.
 BUI86364
 ACCESSION BUI86364.1 GI:22700348
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 896)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: DCTD/DRP/Gardar
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
 http://image.lnl.gov
 Plate: LLM41319 row: b column: 21
 High quality sequence stop: 358.
 Location/Qualifiers
 1. 896
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6010676"
 /tissue_type="large cell carcinoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_1ib="NIH MGC 68"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 1.8 kb. Library constructed by Life
 Technologies."

BASE COUNT 219 a 186 g 264 c
 ORIGIN
 1 227 c 186 g 264 c

Query Match 3.8%; Score 104; DB 13; Length 896;
 Best Local Similarity 100.0%; Pred. No. 5.7e-44;
 Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 TTTCAGATTGTTGAATCTCTGGCCGACACATACAGAGAGACTTAAGCAGCA 177
 Db 327 TTTCAGATTGTTGAATCTCTGGCCGACACATACAGAGAGACTTAAGCAGCA 386
 QY 178 AGGACCTACAGCGTCTGAGCAGATGAGCTGTTACTAGATTG 221
 Db 387 AGGACCTACAGCGTCTGAGCAGATGAGCTGTTACTAGATTG 430

RESULT 23
 BE033028 570 bp mRNA linear EST 09-JUL-2000
 LOCUS 133242 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
 DEFINITION BE033028
 ACCESSION BE033028.1 GI:8327974
 VERSION EST.
 KEYWORDS Sus scrofa (pig)
 SOURCE

ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 1 (bases 1 to 570)
 Fahnenkrug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,
 Vallet, J., Wise, T., Rohrer, G.A., Perlee, G., Sultana, R., Quackenbush,
 J. and Keele, J.W.
 Porcine gene discovery by normalized cDNA-library sequencing and
 EST cluster assembly
 Mamm. Genome 13 (8), 475-478 (2002)
 JOURNAL 22213789
 MEDLINE 12226715
 PUBLISHED
 COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -mismatch 12 options.
 PCR primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCAGTCACGACG
 Plate: 67 row: D column: 12
 Seq primer: ATTAGTGCACATATAG.
 Location/Qualifiers
 1. 570
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_1ib="MARC 1PIG"
 /note="Vector: pCMV SPORT6; Site_1: NotI; Site_2: SalI;
 library made from pooled tissue from day 11, 13, 15, 20,
 and 30 embryos."

BASE COUNT 168 a 109 c 136 g 157 t
 ORIGIN
 1 109 c 136 g 157 t

Query Match 3.7%; Score 101; DB 10; Length 570;
 Best Local Similarity 100.0%; Pred. No. 2.1e-42;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 305 TACAAAGAAATGTGAATCCACATGATGATCACTTCAATGGCTGGCCACAGCTCC 364
 Db 178 TACAAAGAAATGTGAATCCACATGATGATCACTTCAATGGCTGGCCACAGCTCC 237
 QY 365 AGTTATCATACCTCTTTTGGATGAGGACGAGTAGGCT 405
 Db 238 AGTTATCATACCTCTTTTGGATGAGGACGAGTAGGCT 278

RESULT 24
 BM801410 1005 bp mRNA linear EST 05-MAR-2002
 LOCUS AGENCOURT_6459082 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5560757
 DEFINITION 5', mRNA sequence.
 BM801410
 ACCESSION BM801410.1 GI:19118233
 VERSION EST.
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1005)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Unpublished
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.

CDNA library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LLM12287 row: h column: 06
 High quality sequence stop: 500.

FEATURES

source

1. 1005
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5560757"
 /tissue_type="duodenal adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 88"
 /note="Organ: small_intestine; Vector: pCMV-SPORT6;
 site_1: NotI; Site_2: SalI; cloned unidirectionally;
 oligo-dt primed. Average insert size 1.767 kb. Library
 enriched for full-length clones and constructed by Life
 Technologies. Note: this is a NIH_MGC Library."
 BASE COUNT 262 a 283 c 138 g 322 t
 ORIGIN

Query Match 2.5%; Score 68; DB 12; Length 1005;
 Best Local Similarity 100.0%; Pred. No. 1.7e-24;
 Matches 68; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 465 AAAAGATGTCGCGCATATCTTACACGAAAGATGATGACAGTGGCTGGAAAG 524
 |||||
 1 AAAAGATGTCGCGCATATCTTACACGAAAGATGATGACAGTGGCTGGAAAG 60

QY 525 ACATCCTG 532
 |||||
 61 ACATCCTG 68

RESULT 25 AG157495 677 bp DNA linear GSS 09-JAN-2002
 AG157495
 LOCUS
 DEFINITION Pan troglodytes DNA, clone: RP43-022G24.T7, genomic survey
 sequence.
 ACCESSION AG157495
 VERSION AG157495.1 GI:16687173
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL

1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 BAC end sequences of Library RPCI-43
 2 (bases 1 to 677)
 Unpublished
 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,
 Totoki, Y., Watanabe, H. and Sakaki, Y.
 Direct Submission
 Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
 1-7-22 Suenho-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail: chimbases@gsc.riken.go.jp, URL: <http://hgp.gsc.riken.go.jp/>,
 Tel: 81-45-503-9111, Fax: 81-45-503-9170)
 Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
 end was generated during the Rad process and may have higher chance
 of clone tracking errors.
 PRIMERS
 Sequencing: T7
 LIBRARY
 Vector : PBACg3.6
 R.Site 1 : EcoRI
 R.Site 2 : EcoRI.
 Location/Qualifiers

COMMENT
 FEATURES

FEATURES

source

1. 677
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="RP43-022G24.T7"
 /sex="male"
 /cell_type="lymphocytes"
 /clone_lib="RPCI-43 Chimpanzee Male BAC Library"
 BASE COUNT 243 a 106 c 91 g 236 t
 ORIGIN

Query Match 2.5%; Score 67; DB 29; Length 677;
 Best Local Similarity 100.0%; Pred. No. 5.2e-24;
 Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1850 AAGCCACAGACGACAGATATTAAGATGAGACCCAGTGCACGTTCAGACTTA 1909
 |||||
 312 AAGCCACAGACGACAGATATTAAGATGAGACCCAGTGCACGTTCAGACTTA 371

QY 1910 CACCATG 1916
 |||||
 372 CACCATG 378

RESULT 26 B1183160 604 bp mRNA linear EST 10-JUL-2001
 B1183160
 LOCUS
 DEFINITION UNL-P-FN-bl-a-10-0-UNL.s1 UNL-P-FN Sus scrofa cDNA clone
 UNL-P-FN-bl-a-10-0-UNL.3', mRNA sequence.
 ACCESSION B1183160
 VERSION B1183160.1 GI:14657569
 KEYWORDS EST.
 SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED

1 (bases 1 to 604)
 Caetano, A.R., Johnson, R.K. and Pomp, D.
 Generation and sequence characterization of a normalized cDNA
 library from swine ovarian follicles
 Mamm. Genome 14 (1), 65-70 (2003)
 22419904
 12532269
 Contact: Pomp, D
 Department of Animal Science
 University of Nebraska, Lincoln
 Lincoln, NE 68583-0908, USA
 Tel: 402 472 6416
 Fax: 402 472 6362
 Email: dpomp@unl.edu
 Oligo-dt track not found, Not I site shown in beginning of sequence
 is likely internal to the message.
 Seq primer: M13 -29
 POLYA-No.

FEATURES

source

1. 604
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /strain="University of Nebraska, Lincoln Swine Selection
 Lines"
 /db_xref="taxon:9823"
 /clone="UNL-P-FN-bl-a-10-0-UNL"
 /dev_stage="ADULT"
 /lab_host="DH10B (Life Technologies)"
 /clone_lib="UNL-P-FN"
 /note="Vector: pT733-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The UNL-P-FN
 library is a normalized library representing porcine
 ovarian follicles, ranging between 2.0 to 10.0 mm in
 diameter, collected during 7 days of the follicular phase
 of the pig estrous cycle. This library was derived from
 the library UNL-P-F2. The tag is a string of 5-6
 nucleotides present between the Not I site and the

BASE COUNT

180 a _132 c 130 g 162 t

ORIGIN

oligo-dt track. The library was constructed as described by Bonaldo, Lemon and Soares, Genome Research 6: 791-806, 1996.
TAG SEQ=None found"

Query Match 2.2%; Score 59; DB 12; Length 604;
Best Local Similarity 100.0%; Pred. No. 1.1e-19;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 770 GCACGCTGATTTTATGGGCGAGACTTGTCTATCTTCCGAACCTTGGGCGACACCC 828
DB 234 GCACGCTGATTTTATGGGCGAGACTTGTCTATCTTCCGAACCTTGGGCGACACCC 292

RESULT 27

LOCUS

DEFINITION

BX282800 Soares NhhmPu S1 Homo sapiens CDNA clone IMAGE:998N16187 ;
IMAGE:767055, mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

Query Match 2.1%; Score 57; DB 13; Length 598;
Best Local Similarity 100.0%; Pred. No. 1.3e-18;
Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

2522 TTACCTCTAGAACCTCAACAGTAACTGCTAGACATTAACCTGAAACCA 2578
DB 75 TTACCTCTAGAACCTCAACAGTAACTGCTAGACATTAACCTGAAACCA 131

RESULT 28

LOCUS

DEFINITION

AO533647 576 bp DNA linear GSS 18-MAY-1999
RPCT-11-35304.TJ RPCT-11 Homo sapiens genomic clone RPCT-11-35304,
genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

Query Match 2.1%; Score 56; DB 28; Length 576;
Best Local Similarity 100.0%; Pred. No. 4.4e-18;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN

399 GTAGGCTGATTTGAGGAGCAAGATCAATATTTTCATGACCTGGTTAATAC 454
DB 576 GTAGGCTGATTTGAGGAGCAAGATCAATATTTTCATGACCTGGTTAATAC 521

RESULT 29

LOCUS

DEFINITION

CA872840 559 bp mRNA linear EST 20-DEC-2002
K0920H08-5N NIA Mouse Neural Stem Cell (Undifferentiated) cDNA
Library (long) Mus musculus cDNA clone NIA:K0920H08 IMAGE:30086203
5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

COMMENT

ACCESSION CA872840
 VERSION CA872840.1 GI:27324389
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 559)
 AUTHORS Piao, Y., Dudekula, D.B., Qian, Y., Martin, P.R., Alba, K., Vescevi, A.L. and Ko, M.S.H.
 TITLE Systematic Analyses of NIA Mouse Neural Stem Cell (Undifferentiated) cDNA Library (Long)
 JOURNAL Unpublished
 COMMENT Contact: Dawood B. Dudekula
 Laboratory of Genetics
 National Institute on Aging/National Institutes of Health
 333 Caseell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
 Email: cdna@lgsun.grc.nia.nih.gov
 Plate: K0920 row: H column: 08
 Seq primer: -21M13 Reverse
 High quality sequence stop: 559
 POLYVA=No.

FEATURES
 source
 1..559
 /location/Qualifiers
 /organism="Mus musculus"
 /mol_type="CD1"
 /strain="CD1"
 /db_xref="niaEST:K0920H08-5N"
 /db_xref="taxon:10090"
 /clone="NIA:K0920H08 IMAGE:30086203"
 /dev_stage="Adult"
 /lab_host="DHI08"
 /clone_lib="NIA Mouse Neural Stem Cell (Undifferentiated) cDNA Library (Long)"
 /note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA). This is a long-transcript enriched cDNA library (Ref. Genome Res. 11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were obtained from Dr. Angelo L. Vescevi (Institute for Stem Cell Research, Italy). Double-stranded cDNAs were synthesized with an oligo(dT) primer [Invitrogen: 5'-pgACTAGTCTAGATCGCGCGCCGCTTTTCTTTT-3'] from 2.0 Microgram of total RNA, treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to lone-linker lu-Sal4, purified by phenol/chloroform, and separated from free linkers by centrifugion 100. Then, the cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) with a primer Sal4-S. The products were purified by phenol/chloroform and centrifugion 100. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DHI08 B. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is about 3.8 kb. The library was constructed by Yulan Piao."

BASE COUNT 166 a 115 c 126 g 152 t
 ORIGIN

Query Match 1.9%; Score 52; DB 14; Length 559;
 Best Local Similarity 100.0%; Pred. No. 6.3e-16;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 113 GCTATTTCAGATTGTTGAATCTCTCGCGCGCATATACGAGAGAG 164
 |||||
 DB 3 GCTATTTCAGATTGTTGAATCTCTCGCGCGCATATACGAGAGAG 54

RESULT 30
 AA716143 394 bp mRNA linear EST 29-DEC-1997
 LOCUS AA716143/c
 DEFINITION Zg60d02.s1 Soares_fetal_heart_NbH19w Homo sapiens cDNA clone

IMAGE:397731 3', mRNA sequence.
 AA716143.
 AA716143.1 GI:2728417
 EST.
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 394)
 AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Maita, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.
 TITLE Washu-NCI human EST Project
 JOURNAL Unpublished
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watsun.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -40m13 fwd. ET from Amersham
 High quality sequence stop: 363.

FEATURES
 source
 1..394
 /location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="GDB:1305554"
 /db_xref="taxon:9606"
 /clone="IMAGE:397731"
 /sex="unknown"
 /dev_stage="19 weeks"
 /lab_host="DHI08 (ampicillin resistant)"
 /clone_lib="Soares_fetal_heart_NbH19w"
 /note="Organ: heart; Vector: pRTT3D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer (5' TCTTTCAGATCGAAGTGGAGCGCGCGCATATATCTTTTCTTTT 3'), double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pRTT3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M.Fatima Bonaudo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NbH19w."

BASE COUNT 128 a 69 c 62 g 135 t
 ORIGIN

Query Match 1.9%; Score 51; DB 9; Length 394;
 Best Local Similarity 100.0%; Pred. No. 2e-15;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2659 ATAAATTAATCCAGATGATTAATCTCTTAATGCGTTTTCCTAATAC 2709
 |||||
 DB 394 ATAAATTAATCCAGATGATTAATCTCTTAATGCGTTTTCCTAATAC 344

RESULT 31
 BH265538 773 bp DNA linear GSS 30-NOV-2001
 LOCUS BH265538
 DEFINITION CH230-43J18.TU CHORI-230 Segment 1 Rattus norvegicus genomic clone
 CH230-43J18, genomic survey sequence.
 ACCESSION BH265538
 VERSION BH265538.1 GI:17175582
 KEYWORDS GSS.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

REFERENCE 1 (bases 1 to 773)
AUTHORS Zhao, S., Shetty, J., Shateman, S., Tesgaye, G., Geer, K., Shvartsbeyn, A., Gebregeorgis, B., Overton, L., Russell, D., Chen, D., Shyng, F., de Jong, P. and Fraser, C.M.
TITLE Rat BAC End Sequences from Library CHORI-230 EcoRI segment
JOURNAL Unpublished
COMMENT Other GSSs: CH230-43J18.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the rat BAC library CHORI-230
(http://www.chori.org/bacpac/rat230.htm). For BAC library
availability, please contact Pieter de Jong (pdejong@mail.cho.org).
Clones may be purchased from BACPAC Resources
(http://www.chori.org/bacpac/orering/information.htm). BAC end
page: http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html
Plate: 43 row: J column: 18
Seq primer: SP6
Class: BAC ends.

FEATURES
source
1. .773
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/NaNHed/MCW"
/db_xref="taxon:10116"
/clone="CH230-43J18"
/sex="Female"
/cell_type="Brain"
/note="Vector: pTARBAC2.1; Site 1: EcoRI; Site 2: EcoRI;
CHORI-230 Rat (BN/NaNHed/MCW) BAC library produced by
Pieter de Jong"

BASE COUNT 240 a 159 c 157 g 217 t

ORIGIN

Query Match 1.8%; Score 50; DB 28; Length 773;
Best Local Similarity 100.0%; Pred.No. 8.3e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2291 AGCCAGAGGTCTGTACAGAGACTTCATGACCTCATCAACCCCA 2340
|||||
571 AGCCAGAGGTCTGTACAGAGACTTCATGACCTCATCAACCCCA 620

Db

RESULT 32
BG315503 124 bp mRNA linear EST 28-FEB-2002
LOCUS POS 0.131 Human THP1 cell line library Homo sapiens cDNA, mRNA
DEFINITION
ACCESSION BG315503
VERSION BG315503.1 GI:18998375
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 124)
Andersson, T., Borzang, S., Larsson, M., Thelin, A., Ekstrand-Hammarstrom
B., Witta, V., Wenborg, A., Lundberg, J. and Odeberg, J.
Identification of candidate genes in atherosclerosis - Virtual chip
analysis in RDA based transcript profiling of monocyte/macrophage
response to oxidised LDL
Unpublished
Contact: Andersson Tove
Department of Biotechnology
KTH
Teknikringen 34, plan 6, 100 44 Stockholm, Sweden
Tel: +46 8 790 71 29
Fax: +46 8 245452

Email: tove@biochem.kth.se
POLYA-No.

FEATURES
source
1.124
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cell_line="THP1"
/note="Vector: pRIT28; Site 1: BamHI; Site 2: BamHI;
Shotgun cloning of RDA difference products. Macrophage and
foamcell libraries were submitted to successive rounds of
subtractive hybridizations generating populations of gene
fragments that are differentially expressed in macrophage
to foam cell formation."

BASE COUNT 37 a 18 c 29 g 38 t 2 others

ORIGIN

Query Match 1.8%; Score 49; DB 10; Length 124;
Best Local Similarity 100.0%; Pred.No. 1.8e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1551 TCGAAGAGATGTTGGACCGCTTCTTAAGATGTTCAATCTTAAGA 1599
|||||
76 TCGAAGAGATGTTGGACCGCTTCTTAAGATGTTCAATCTTAAGA 124

Db

RESULT 33
BY021039 371 bp mRNA linear EST 06-DEC-2002
LOCUS BY021039 RIKEN full-length enriched, mammary gland RCB-0526
DEFINITION JY9-MC(A) cDNA Mus musculus cDNA clone G830044019 5', mRNA
sequence.
ACCESSION BY021039
VERSION BY021039.1 GI:26126482
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 371)
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H.,
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
Gotohori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A.,
Quackenbush, J., Schiml, L.M., Kanapin, A., Matsuda, H., Batalov, S.,
Beisel, K.W., Blake, J.A., Brad, D., Brusic, V., Chochia, C., Corbani,
L.E., Cousins, S., Dalla, E., Dregani, T.A., Fletcher, C.F., Forrest,
A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A.,
Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J.,
Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M.,
King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons,
P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki,
H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Pertea, G.,
Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.V., Qi, D.,
Ramachandran, S., Ravasi, T., Reed, J.C., Reed, J.U., Reid, J., Ring,
B.Z., Ringwald, M., Sanderlin, A., Schneider, C., Semple, C.A., Setou,
M., Shimada, K., Sultana, R., Takanaka, Y., Taylor, M.S., Teasdale,
R.D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L.G., Wyshaw-Borja, A., Yanagisawa,
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
Carrinzi, P., Hayatsu, N., Hirozane-Kikukawa, T., Kono, H., Nakamura,
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii,
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
E.S., Rogers, J., Birney, E., Birney, E., and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)
MEDLINE 22354683
PUBMED 12466851
COMMENT Contact: Yoshinori Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gscc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watabiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

```
1..371
/location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="G830044J19"
/tissue_type="mammary gland"
/cell_line="RCB-0526 Jyg-MC(A)"
/clone_lib="RIKEN full-length enriched, mammary gland
RCB-0526 Jyg-MC(A) cDNA"
```

BASE COUNT

```
96 a 95 c 84 g 96 t
```

Query Match

```
Best Local Similarity 1.7%; Score 47; DB 13; Length 371;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
118 TTTCAGATTGTTGAACCTTCTCTGCGCAGCATACAGAGAGAG 164
104 TTTCAGATTGTTGAACCTTCTCTGCGCAGCATACAGAGAGAG 150
```

DB

```
RESULT 34
BY018262 401 bp mRNA linear EST 06-DEC-2002
```

```
LOCUS BY018262
DEFINITION BY018262 RIKEN full-length enriched, mammary gland RCB-0526
Jyg-MC(A) cDNA Mus musculus cDNA clone G830027G18 5', mRNA
sequence.
```

```
ACCESSION BY018262
VERSION BY018262.1 GI:26078511
KEYWORDS EST
```

```
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
```

```
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclerognathi; Muridae; Murinae; Mus.
1 (bases 1 to 401)
```

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,
Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamana, K., Kiyosawa, H.,
Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,
Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C., Hume, D.A.,
Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batalov, S.,

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

Beisel, K.W., Blake, J.A., Brad, D., Brusic, V., Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimond, S., Gustincich, S., Hirokawa, N., Jackson, I.O., Jarvis, E.D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavani, W.J., Pereira, G., Penelope, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, U., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Vezardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kikukawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shingawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

CONTACT

Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@gscc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watabiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

source

```
1..401
/location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="G830027G18"
/tissue_type="mammary gland"
/cell_line="RCB-0526 Jyg-MC(A)"
/clone_lib="RIKEN full-length enriched, mammary gland
RCB-0526 Jyg-MC(A) cDNA"
```

BASE COUNT

```
107 a 97 c 94 g 102 t 1 others
```

Query Match

```
1.7%; Score 47; DB 13; Length 401;
```

Best Local Similarity 100.0%; Pred. No. 2,9e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 TTTCAGATTGTTGAACCTCTCTGCGCACAATACGAGAGGAG 164
|||||
Db 104 TTTCAGATTGTTGAACCTCTCTGCGCACAATACGAGAGGAG 150

RESULT 35

BY261117

LOCUS BY261117 RIKEN full-length enriched, visual cortex Mus musculus 450 bp mRNA linear EST 10-DEC-2002

DEFINITION cDNA clone K330305G13 5', mRNA sequence.

ACCESSION

BY261117

VERSION BY261117.1 GI:26442629

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

AUTHORS

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Oosato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schombach, C., Gojbori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Choitha, C., Corbett, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Fraser, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jaray, B. D., Kana, A., Kawai, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lemhard, B., Lyons, P. A., Maglott, D. R., Malais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Petosa, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J. C., Reed, J. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Sempke, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wymshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavalon, M., Zhu, Y., Zimmer, A., Carrinci, P., Hayatsu, N., Hirozane-Kishikawa, T., Komno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Akawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

TITLE

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL

NATURE 420, 563-573 (2002)

MEDLINE

22354683

PUBMED

12466851

COMMENT

Contact: Yoshitake Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp/
URL: http://genome.gsc.riken.go.jp/
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Komno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tegan, M., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y. Direct Substitution
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1157-1171 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Exploration Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
Tissues were provided by Michela Fagioli and Takao K. Henrich (Laboratory for Neuronal Circuit Development Brain Science Institute RIKEN 2-1 Hirosewa, Wako-shi, Saitama 351-0198 Japan) whose assistance we gratefully acknowledge. Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.
location/Qualifiers

FEATURES

source

1.450
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="K330305G13"
/tissue_type="visual cortex"
/clone_lib="RIKEN full-length enriched, visual cortex"
130 a 96 c 102 g 122 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 3e-13;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 TTTCAGATTGTTGAACCTCTCTGCGCACAATACGAGAGGAG 164
|||||
Db 95 TTTCAGATTGTTGAACCTCTCTGCGCACAATACGAGAGGAG 141

RESULT 36

BI465498

LOCUS BI465498 470 bp mRNA linear EST 22-AUG-2001

DEFINITION field1.yl Kaestner ngm wt Mus musculus cDNA 5' similar to TR:063548 Q63548 COLAPSN-1 PRECURSOR, mRNA sequence.

ACCESSION BI465498

VERSION BI465498.1 GI:15278199

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 470)

AUTHORS Melton, D., Brown, J., Kently, G., Permut, A., Lee, C., Kaestner, K., Lemishka, I., Scarce, M., Bresnelli, J., Gradwohl, G., Clifton, S., Hillier, L., Maira, M., Page, D., Wylie, T., Martin, J., Bristain, A., Schmitt, A., Theising, B., Rittler, B., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagaris, R., Williams, T., Jackson, Y. and Bowers, Y.

TITLE Endocrine Pancreas Consortium Unpublished

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge, MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@oeb.harvard.edu

Pancreas was obtained from Gerard Gradwohl (PNAS 97 P1607-1611, 2000) Library was constructed by Catherine Lee DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Dr. Marie Scarce (mscarce@mail.med.upenn.edu)

High quality sequence stop: 348.

Location/Qualifiers

1.470

FEATURES

source

```

/organism="Mus musculus"
/mol_type="mRNA"
/strain="129/Sv x CD1"
/db_xref="taxon:10090"
/dev_stage="p.c.14.5"
/lab_host="E. coli DH2S (GIBCO)"
/clone_lib="Kaesler ngn3 wt"
/notes="Organ: pancreas; Vector: pSPORT1 (GIBCO); Site_1: Not 1; Site_2: Sal I; The library was prepared by Catherine S. Lee and has not been published. The pancreas was obtained from Gerard Gradwohl (PNAS 97 p1607-1611, 2000). The cDNAs were prepared with an oligo containing a NotI site, and SalI linkers were added to the ends. The inserts were cut with NotI before being cloned into the NotI-SalI sites in the vectors. This is one of two libraries, ngn3 wt and ngn3 -/- . The wt library is in pSPORT1, T7 promoter is 5'."

BASE COUNT      133 a      97 c      106 g      134 t
ORIGIN

Query Match      1.7%; Score 47; DB 12; Length 470;
Best Local Similarity 100.0%; Pred. No. 3e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

118 TTTCAGATTGTTGAATCTCTCTGCGCCGACATACAGAGAAG 164
|||||
116 TTTCAGATTGTTGAATCTCTCTGCGCCGACATACAGAGAAG 162

RESULT 37
LOCUS      BB655104      639 bp      mRNA      linear      EST 26-OCT-2001
DEFINITION BB655104 RIKEN full-length enriched, 9 days embryo Mus musculus
CDNA clone D030032A02 5', mRNA sequence.
BB655104
VERSION     BB655104.1 GI:16488932
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 639)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel.: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-ree@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
Wagai,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watanabe,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura
,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara

```

```

,Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nondundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanka,I., Aizawa
,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and
Hayashizaki,Y.
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
Location/Qualifiers
1..639
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone_lib="D030032A02"
/dev_stage="9 days embryo"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 9 days embryo"
/notes="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer 15'
GAGGAGAGAGCGCGCCGCACTCGAGTTTCTTTTCTTTTCTTTT 3'1. cDNA was
prepared by using trihaloalkano thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence 15'
GAGGAGAGAGTTCTCGAGTTTCTTTTCTTTTCTTTTCTTTT 3'1. cDNA
was cleaved with BamI and XhoI. Vector: a modified
pBluescript KS(+) after bulk excision from lambda FLIC I."

BASE COUNT      183 a      147 c      155 g      154 t
ORIGIN

Query Match      1.7%; Score 47; DB 10; Length 639;
Best Local Similarity 100.0%; Pred. No. 3.3e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

118 TTTCAGATTGTTGAATCTCTCTGCGCCGACATACAGAGAAG 164
|||||
343 TTTCAGATTGTTGAATCTCTCTGCGCCGACATACAGAGAAG 389

RESULT 38
LOCUS      BB663470      653 bp      mRNA      linear      EST 26-OCT-2001
DEFINITION BB663470 RIKEN full-length enriched, 15 days embryo head Mus
musculus cDNA clone D930045G07 5', mRNA sequence.
BB663470
VERSION     BB663470.1 GI:16497224
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 653)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki
,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute

```

The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura
, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Konno, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa
, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
Hayashizaki, Y.
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.

FEATURES

source

```
1. 653
Location/Qualifiers
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="D930045G07"
/sex="mixed"
/tissue_type="head"
/dev_stage="15 days embryo"
/lab_host="DH10B"
/clone_1ib="RIKEN full-length enriched, 15 days embryo
head"
/note="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGATTCGACGCTCTTTTCTTTTCTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGAGAGATTCGACGCTCTTTTCTTTTCTTTTNN 3']. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I"
```

BASE COUNT

176 a 149 c 161 g 167 t

Query Match 1.7%; Score 47; DB 10; Length 653;

Best Local Similarity 100.0%; Pred. No. 3-3e-13; Mismatches 0; Indels 0; Gaps 0;

118 TTCCGATTTGTTGAATCTCTGCGCCGACATACAGAGAGAG 164
DB 409 TTCCGATTTGTTGAATCTCTGCGCCGACATACAGAGAG 455

RESULT 39

BB663737
LOCUS BB663737 655 bp mRNA linear EST 26-OCT-2001
DEFINITION BB663737 RIKEN full-length enriched, 0 day neonate lung Mus
musculus cDNA clone E030004A12 5', mRNA sequence.
ACCESSION BB663737
VERSION BB663737.1 GI:16497491
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Unpublished
Contract: Yoshida Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Wataniki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunura
, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Konno, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa
, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
Hayashizaki, Y.
Computational Analysis of Full-length Mouse cDNAs Compared with
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.

FEATURES

source

```
1. 655
Location/Qualifiers
/mol_type="mRNA"
/strain="E030004A12"
/db_xref="taxon:10090"
/clone="E030004A12"
/tissue_type="lung"
/dev_stage="0 day neonate"
/lab_host="DH10B"
/clone_1ib="RIKEN full-length enriched, 0 day neonate
lung"
/note="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
```

GAGAGAGAGAGCGCGCCGCACTCGAGTTTCTTTTCTTTT 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGAGAGAGATTCGAGTTTAATTAATTAATTCCTCCCTCCCTCC 3'. cDNA was cleaved with BamHI and XhoI. Vector: a modified pluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 157 a 140 c 125 g 233 t

Query Match 1.7%; Score 47; DB 10; Length 655;
Best Local Similarity 100.0%; Pred. No. 3.3e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 118 TTTCAGATTGTTGAATCTCTGCGCCGCAATACAGAGAGAG 164
Db TTTCAGATTGTTGAATCTCTGCGCCGCAATACAGAGAGAG 344

RESULT 40
LOCUS BB658393 657 bp mRNA linear EST 26-OCT-2001
DEFINITION BB658393 RIKEN full-length enriched, 12 days embryo eyeball Mus musculus cDNA clone D230041C01 5', mRNA sequence.

ACCESSION BB658393
VERSION BB658393.1 GI:16492218
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 657)
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Komoto, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.go.jp,
URL: http://genome-gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komoto, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
Wagii, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, B., Mochizuki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, Y. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamataka, I., Adzawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)

Please visit our web site (<http://genome-gsc.riken.go.jp>) for further details.
e mouse tissues.
location/Qualifiers

FEATURES
source
1. 657
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="D230041C01"
/tissue_type="eyeball"
/dev_stage="12 days embryo"
/lab_host="DH10B"
/clone_id="RIKEN full-length enriched, 12 days embryo eyeball"
/note="Site 1: SalI; Site 2: BamHI. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer 15' GAGAGAGAGCGCGCCGCACTCGAGTTTCTTTTCTTTT 3'. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence 15' GAGAGAGATTCGAGTTTAATTAATTAATTCCTCCCTCCCTCC 3'. cDNA was cleaved with BamHI and XhoI. Vector: a modified pluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 199 a 144 c 154 g 160 t

Query Match 1.7%; Score 47; DB 10; Length 657;
Best Local Similarity 100.0%; Pred. No. 3.3e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 118 TTTCAGATTGTTGAATCTCTGCGCCGCAATACAGAGAGAG 164
Db TTTCAGATTGTTGAATCTCTGCGCCGCAATACAGAGAGAG 364

RESULT 41
LOCUS BY734980 677 bp mRNA linear EST 17-DEC-2002
DEFINITION BY734980 RIKEN full-length enriched, mammary gland RCB-0526 Uyg-MC(A) cDNA Mus musculus cDNA clone G830046C02 5', mRNA sequence.

ACCESSION BY734980
VERSION BY734980.1 GI:27148107
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 677)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nishida, I., Ose, N., Saito, R., Suzuki, H., Yamataka, I., Kiyosawa, H., Yagii, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bull, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matuda, H., Batalov, S., Betts, K. W., Blake, J. A., Brad, D., Brusic, V., Chochia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pereira, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sanderlin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verrardo, R., Wagner, L., Wahlestedt, C., Wang, Y.,

Watanabe, Y., Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, I., Aizawa, K., Aikawa, T., Fukuda, S., Hara, A., Hashizume, W., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Yoshizaki, Y.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)

JOURNAL
 MEDLINE
 PUBLISHED
 22354683
 12466851

COMMENT
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gs.c.riken.go.jp/
 URL: http://genome.gsc.riken.go.jp/
 Adachi, J., Aizawa, K., Akimura, T., Aikawa, T., Carninci, P., Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Ohnato, N., Saito, R., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Takeda, Y., Waki, K., Watabiki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site (http://genome.gsc.riken.go.jp) for further details.

FEATURES

source

Location/Qualifiers
 1. 677
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="G830046C02"
 /tissue_type="mammary gland"
 /cell_line="RCB-0526 Jyg-MC(A)"
 /clone_1ib="RIKEN full-length enriched, mammary gland RCB-0526 Jyg-MC(A) cDNA"
 BASE COUNT 130 a 167 c 149 g 224 t 7 others

Query Match 1.7%; Score 47; DB 14; Length 677;
 Best Local Similarity 100.0%; Pred. No. 3-3e-13;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 TTTCAGATTGTTGAATCTCTGCGCCGACAAATACAGAGAG 164
 |||||
 DB 478 TTTCAGATTGTTGAATCTCTGCGCCGACAAATACAGAGAG 524

RESULT 42
 BU708579 731 bp mRNA linear EST 26-NOV-2002
 LOCUS BU708579

DEFINITION
 UI-M-F10-cat-d-15-0-UI-r1 NIH BMAP F10 Mus musculus cDNA clone
 IMAGE: 6416990 5', mRNA sequence.
 BU708579
 BU708579.1 GI:23641181
 EST.
 KEYWORDS
 Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 731)
 NIH-MGC http://mgc.nci.nih.gov/.
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL
 TITLE
 Unpublished
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bcr-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
 The following repetitive elements were found in this cDNA sequence: 35-96, >POLY_A#Simple_repeat (matched complement)
 Seq primer: pYX-5

FEATURES

source

Location/Qualifiers
 1. 731
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE: 6416990"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_1ib="NIH BMAP F10"
 /note="Organ: Brain; Vector: pYX-Abs; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lemmon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to RNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Abs vector. The library tag sequence located between the Not I site and the polyA tail is CAGCCAGCAGC. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemlin Chin, Ph.D., program coordinator."

BASE COUNT 187 a 146 c 156 g 242 t

Query Match 1.7%; Score 47; DB 13; Length 731;
 Best Local Similarity 100.0%; Pred. No. 3-4e-13;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 TTTCAGATTGTTGAATCTCTGCGCCGACAAATACAGAGAG 164
 |||||
 DB 215 TTTCAGATTGTTGAATCTCTGCGCCGACAAATACAGAGAG 261

RESULT 43
 CB518391 787 bp mRNA linear EST 28-MAR-2003
 LOCUS CB518391
 DEFINITION UI-M-GHO-csc-0-11-0-UI-r1 NIH BMAP_GHO Mus musculus cDNA clone
 IMAGE: 6837828 5', mRNA sequence.
 ACCESSION CB518391
 VERSION CB518391.1 GI:29351746
 KEYWORDS EST.

SOURCE
ORGANISM Mus musculus (house mouse)
Mus musculus
REFERENCE Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 787)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bhs-remail.nih.gov
 Tissue Procurement: Dr. Jim Lin, University of Iowa
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at <http://genome.uiowa.edu/distribution/mouseefl.html>
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP)
 The following repetitive elements were found in this cDNA sequence: 94-126, >(GAA)n#simple repeat (matched complement) 169-233, >POLY A#simple_repeat (matched complement)
 Seq primer: pyx-5

FEATURES
source
 Location/Qualifiers
 1..787
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE: 6837828"
 /cissue_type="Whole brain"
 /dev_stage="1, 5, and 15 days newborn"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP GH0"
 /note="Organ: Brain; Vector: pyx-Asc; Site:1: EcoR I; Site:2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pyx-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAAGTGAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., Program coordinator."

BASE COUNT
 200 a 160 c 170 g 256 t 1 others

ORIGIN

Query Match 1.7%; Score 47; DB 14; Length 787;
 Best Local Similarity 100.0%; Pred. No. 3.5e-13;
 Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 TTTCAGATTTGTTGAATCTCTGCGCACATACAGAGAAG-164
 |||||
 352 TTTCAGATTTGTTGAATCTCTGCGCACATACAGAGAAG-398

Db 352 TTTCAGATTTGTTGAATCTCTGCGCACATACAGAGAAG-398

RESULT 44
LOCUS AK053115 2596 bp mRNA linear HTC 05-DEC-2002
DEFINITION Mus musculus 0 day neonate lung cDNA, RIKEN full-length enriched library, clone: E030004A12 product: sem domain, immunoglobulin domain (Ig), short basic domain, secreted, (semaphorin) 3A, full insert sequence.
ACCESSION AK053115
VERSION AK053115.1 GI:26343226
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS Carninci, P., and Hayashizaki, Y.
TITLE High-efficiency full-length cDNA cloning
JOURNAL Meth. Enzymol. 303, 19-44 (1999)
MEDLINE 99279253
PUBMED 10349636
AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL Genome Res. 10 (10), 1617-1630 (2000)
MEDLINE 20493774
PUBMED 11042159
AUTHORS Shibata, K., Itoh, M., Aizawa, K., Nagaoaka, S., Saeki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.
REFERENCE RIKEN integrated sequence analysis (RISA) system-384-Format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
TITLE 20530913
JOURNAL 11076861
MEDLINE 11076861
PUBMED 11076861
AUTHORS Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Aizawa, T., Hara, A., Fukunishi, Y., Kono, H., Adachi, J., Fukuda, S., Arizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamataka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nkaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stabli, F., Suzuki, R., Tomita, M., Wagner, J., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M. F., Brownstein, M. J., Bull, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombauts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyo-oka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohzuki, S., and Hayashizaki, Y.
TITLE Functional annotation of a full-length mouse cDNA collection
JOURNAL Nature 409 (6821), 665-690 (2001)
MEDLINE 21085660
PUBMED 11217851
AUTHORS 5

TITLE The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL Nature 420, 563-573 (2002)
MEDLINE 12108560
PUBMED 12108560
AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arawaka, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Hori, F., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanakawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takahashi, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.
TITLE Direct Submission
 Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of

COMMENT

Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Please visit our web site for further details.
URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>.

FEATURES

source

Location/Qualifiers

1.2596
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM:DB:E030004A12"
/db_xref="taxon:10090"
/clone="E030004A12"
/tissue_type="lung"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="0 day neonate"
383..895
/note="unnamed protein product; putative
sema domain, immunoglobulin domain (Ig), short basic
domain, secreted, (semaphorin) 3A (MGI:107558,
GB|NM_009152, evidence: BLASTN, 99%, match=564)"
/codon_start=1
/protein_id="BAC35270.1"
/db_xref="GI:26343227"
/translation="MGMPGCIACLFNGCVLLTAPANYANGNNVPLKLSYKMLESN
VITFNGLANSSYHTFLIDBERSLVYGAQDHPFRLVNIKQFQITWPSVSTRDE
CKWAGDILKECANFLKVLAVNYQTHLVAGTAPHPICLYIEVGHPEAVICCTELM
LSFGACEGTT"
2575..2580
/note="putative"
2596
/note="putative"

CDS

BASE COUNT 822 a 475 c 541 g 758 t
ORIGIN
polyA_signal
polyA_site
2596
/note="putative"
2596
/note="putative"

Query Match

Best Local Similarity 100.0%; Pred. No. 4,8e-13; Length 2596;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

118 TTTCAGATTGTTGAACCTCTCTGCGCCGACATACAGAGAGAG 164

301 TTTCAGATTGTTGAACCTCTCTGCGCCGACATACAGAGAGAG 347

RESULT 45

BY330392 318 bp mRNA linear EST 11-DEC-2002

LOCUS BY330392 RIKEN full-length enriched, synovial fibroblasts Mus

DEFINITION Musculus cDNA clone LI30014K19 5', mRNA sequence.

ACCESSION BY330392

VERSION BY330392.1 GI:26521121

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 318)

Okaaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I., Kiyosawa, H.,

Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C.,

Gojobori, T., Baldarelli, R., Hill, D. P., Buit, C., Hume, D. A.,

Quackenbush, J., Schriml, J. W., Kanapin, A., Matsuda, H., Batalov, S.,

Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V., Chothia, C., Corbani,

L'E., Cousins, S., Dalla, E., Dregani, T. A., Fletcher, C. F., Forrest

TITLE

JOURNAL
MEDLINE
PUBMED

COMMENT

A., Frazer, K. S., Gaasterland, T., Gariboldi, M., Glisic, C., Godzik, A.,
Gough, J., Grimond, S., Gustincich, S., Hixkawa, N., Jackson, I. J.,
Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedierski, R. M.,
King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons,
H. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, J., Miki,
H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Pertea, G.,
Pesole, G., Petrovsky, N., Pilla, R., Pontus, J. U., Qi, D., Ring,
Ramchandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring,
B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou,
M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale,
R. D., Tomita, M., Verardo, R., Wagner, L., Wahlstedt, C., Wang, Y.,
Watanabe, Y., Wells, C., Wilming, L. G., Wyszewski, A., Yanagisawa,
M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A.,
Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura,
M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, M., Imotani, K., Ishii,
Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata,
K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander,
E. S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gscc.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane,
T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A.,
Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K.,
Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami,
M., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y. Direct
Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.

Tissues were provided by Vassilis Aidinis (Biomedical Sciences
Research Center 'Al. Fleming' Institute of Immunology 14-16 Al.
Fleming street 16672 Vari, Greece) whose assistance we gratefully
acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.

FEATURES

source

Location/Qualifiers

1..318
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="LI30014K19"
/cell_type="synovial fibroblasts"
/clone_lib="RIKEN full-length enriched, synovial
fibroblasts"
BASE COUNT 92 a 72 c 75 g 79 t
ORIGIN

Query Match 1.5%; Score 41; DB 13; Length 318;
Best Local Similarity 100.0%; Pred. No. 4,8e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 124 GATTGTTGAACCTCTCTGCGCCGACAAATACAGACGAG 164
DB 1 GATTGTTGAACCTCTCTGCGCCGACAAATACAGACGAG 41

RESULT 46
CB595903 807 bp mRNA linear EST 03-APR-2003

LOCUS
DEFINITION AGENCOURT_13020395 NIH_MGC_178 Mus musculus cDNA clone
IMAGE:30299607 5', mRNA sequence.

ACCESSION
CB595903
VERSION
CB595903.1 GI:29513759

KEYWORDS
EST.
SOURCE Mus musculus (house mouse)

ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 807)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
CDNA Library Preparation: Michael Brownstein Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: NDCM63 row: p column: 16
High quality sequence stop: 471.

Location/Qualifiers
1. 807

FEATURES
source
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30299607"
/lab_host="DH10B (T1-Phage-resistant)"
/clone_id="NIH_MGC_178"
/note="Organ: lung and heart; Vector: pDNR-LIB; Site 1:
SfiI (ggccattatggcc); Site 2: SfiI (ggccgctggcc); cDNA
made by oligo-dt priming and directionally cloned. 5' and
3' adaptors were used in cloning as follows:
5'-AAGCAGTGTATCAACGAGAGTGGCATTACGCGCG-3' and
5'-ATTCTAGAGCGCGAGCGCGCATG-dt(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the 0.5 kb
size fraction. Library created in the laboratory of M.
Brownstein (NIH). Note: this is a NIH_MGC Library."

BASE COUNT 260 a 175 c 204 g 168 t
ORIGIN

Query Match 1.5%; Score 41; DB 14; Length 807;
Best Local Similarity 100.0%; Pred. No. 6.1e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2171 CAACTCTTCTTAAGTAACTCGAAGTCATTGACACAGA 2211
DB 240 CAACTCTTCTTAAGTAACTCGAAGTCATTGACACAGA 280

RESULT 47

CB196368

LOCUS

DEFINITION

IMAGE:30138167 5', mRNA sequence.

ACCESSION

CB196368

VERSION

900 bp mRNA linear EST 05-FEB-2003
AGENCOURT_11259678 NIH_MGC_135 Mus musculus cDNA clone
IMAGE:30138167 5', mRNA sequence.
CB196368.1 GI:28224005

KEYWORDS
EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
NIH-MGC http://mgi.nci.nih.gov/
1 (bases 1 to 900)
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Dr. David Rowe
CDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: NDM0040 row: 1 column: 24
High quality sequence stop: 655.

Location/Qualifiers
1. 900

FEATURES
source

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30138167"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_135"
/note="Vector: pCMVSPORT1; Site 1: EcoRV; Site 2: NotI;
Normalized full-length enriched library from pooled mouse
embryonic limb, maxilla and mandible, day 12.5, 13.5, 14.5
and 15.5 (size selected for the 0.5-1 kb fragments)
Cloned directionally, priming method: Oligo-dt. cDNA
enrichment: >1k bp. Average insert size 1.6k bp.
Normalization (Cot value): 7.5 kb. Priming sequence:
5'-GACTAGTCTTGAATCGGAGCGCGCCCTT-3' Tissue contributed by
David Rowe. Library constructed by Resgen, Invitrogen
Corp."

BASE COUNT 286 a 208 c 225 g 178 t 3 others
ORIGIN

Query Match 1.5%; Score 41; DB 14; Length 900;
Best Local Similarity 100.0%; Pred. No. 6.3e-10;
Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2171 CAACTCTTCTTAAGTAACTCGAAGTCATTGACACAGA 2211
DB 475 CAACTCTTCTTAAGTAACTCGAAGTCATTGACACAGA 515

RESULT 48

B0884924

LOCUS

DEFINITION

AGENCOURT_8750938 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6336043

VERSION

B0884924

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

NIH-MGC http://mgi.nci.nih.gov/

1 (bases 1 to 917)

Tissue Procurement: Mark Macconochie, Ph.D. and Nancy L. Freeman,

Ph.D.

CDNA Library Preparation: Resgen, Invitrogen Corp

DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: <http://image.liml.gov>
 plate: L1AM13799 row: g column: 20
 High quality sequence stop: 724.
 Location/Qualifiers

FEATURES

source

1. .917
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:6336043"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_130"
 /note="Organ: oocytes; Vector: pCMV-SPORT6.1.ccdp;
 Site 1: EcoRV; Site 2: NotI; Cloned unidirectionally.
 Primer: Oligo dt. Average insert size 1.95 kb.
 Constructed by Resgen, Invitrogen Corp. Note: this is a
 NIH_MGC library."

BASE COUNT 302 a 192 c 217 g 206 t

ORIGIN

Query Match 1.5%; Score 41; DB 13; Length 917;
 Best Local Similarity 100.0%; Pred. No. 6.3e-10;
 Matches 41; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2171 CAACCTCTTCTTAAGTAACTGGAAGTCACTTGACACAGA 2211

DB 39 CAACCTCTTCTTAAGTAACTGGAAGTCACTTGACACAGA 79

RESULT 49
 CB776561 394 bp mRNA linear EST 16-MAY-2003
 LOCUS AMGNNUC:SRP92-00038-g4-A SRP92 (10238) Rattus norvegicus cDNA clone
 DEFINITION SRP92-00038-g4 5', mRNA sequence.
 ACCESSION CB776561 GI:29864952
 VERSION CB776561.1 GI:29864952
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 394)
 Amgen EST Program.
 Amgen Rat EST Program
 JOURNAL Unpublished
 COMMENT Contact: Dan Fitzpatrick
 Amgen, Inc
 One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
 Tel: 805 447-4881
 Plate: 00038 row: g column: 4.

FEATURES

source

1. .394
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="srp92-00038-g4"
 /issue_type="peneal gland brain"
 /clone_lib="srp92 (10238)"
 /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; peneal
 gland brain region"

BASE COUNT 127 a 88 c 78 g 101 t

ORIGIN

Query Match 1.4%; Score 38; DB 14; Length 394;
 Best Local Similarity 100.0%; Pred. No. 2.1e-08;
 Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 986 GCTAGATAGTGTAGATATGACAGATGACTTGAGAG 1023

DB 168 GCTAGATAGTGTAGATATGACAGATGACTTGAGAG 205

RESULT 50
 B0779621/c 657 bp mRNA linear EST 26-JUL-2002
 LOCUS UI-R-PF0-cox-f-23-0-UI.s1 UI-R-PF0 Rattus norvegicus cDNA clone
 DEFINITION UI-R-PF0-cox-f-23-0-UI 3', mRNA sequence.
 ACCESSION B0779621 GI:21988093
 VERSION B0779621
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 657)
 Ronaldo, M.F., Lennon, G. and Soares, M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 JOURNAL
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@iowa.edu
 Tissue Procurement: Jeff Stevens
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: DISTRIBUTION: Researchers may obtain clones
 from Research Genetics (www.resgen.com).
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES

source

Location/Qualifiers
 1. .657
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="UI-R-PF0-cox-f-23-0-UI"
 /issue_type="Mixed tissues"
 /dev_stage="Adult"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /clone_lib="UI-R-PF0"
 /note="Vector: pRT3-Pac (Pharmacia) with a modified
 polylinker; Site_1: EcoR I; Site_2: Not I; UI-R-PF0 is a
 subtracted cDNA library containing the following tissue(s)
): Normal cartilage and SR-JWS Tumor line. The
 subtraction was made according to Ronaldo, Lennon and
 Soares, Genome Research, 6:791-806, 1996. The
 oligonucleotide used to prime the synthesis of
 first-strand cDNA contains a library tag sequence that is
 located between the Not I site and the (dt)18 tail. The
 sequence tags for these libraries are: CTATGAGAG,
 CATCTGTGTA.
 TAG LIB=UI-R-PF0
 TAG_TISSUE=cartilage
 TAG_SEQ=CTATGAGAG"

BASE COUNT 182 a 131 c 117 g 227 t

ORIGIN

Query Match 1.4%; Score 37; DB 13; Length 657;
 Best Local Similarity 100.0%; Pred. No. 8.4e-08;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2600 TTTCATGCGATTATGATGATGTTTCAATGATGGGAA 2636

DB 600 TTTCATGCGATTATGATGATGTTTCAATGATGGGAA 564

RESULT 51

BUT58296/c
 LOCUS BUT58296 658 bp mRNA linear EST 10-OCT-2002
 DEFINITION UI-R-FFO-cov-a-17-0-UI.s1 UI-R-FFO Rattus norvegicus cDNA clone
 ACCESSION UI-R-FFO-cov-a-17-0-UI 3', mRNA sequence.
 VERSION BUT58296
 KEYWORDS BUT58296.1 GI:23720667
 SOURCE EST.
 ORGANISM Rattus norvegicus (Norway rat)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 REFERENCE 1 (bases 1 to 658)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 PUBMED 8889548
 COMMENT Contact: Soares, MB
 Coordinated Laboratory for Computational Genomics
 University of Iowa
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
 Tel: 319 335 8250
 Fax: 319 335 9565
 Email: bento-soares@iowa.edu
 Tissue Procurement: Jeff Stevens, University of Iowa
 cDNA library preparation: Dr. M. Bento Soares, University of Iowa
 cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: DISTRIBUTION: Researchers may obtain clones from Research Genetics (www.reagen.com).
 Seq primer: MJ3 FORWARD
 POLYA=Yes
 FEATURES
 source
 Location/Qualifiers
 1..658
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="UI-R-FFO-cov-a-17-0-UI"
 /issue_type="Mixed tissues"
 /dev_stage="Adult"
 /lab_host="UI-R-FFO"
 /clone_lib="UI-R-FFO"
 /note="Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-R-FFO is a subtracted cDNA library containing the following tissue(s): Normal cartilage and SR-JMS Tumor line. The subtraction was made according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dt)18 tail. The sequence tags for these libraries are: CTATGACAC, CATCTCTGA.
 TAG LIB=UI-R-FFO
 TAG TISSUE=cartilage
 TAG_SEQ=CTATGACG
 BASE COUNT 181 a 131 c 117 g 227 t 2 others
 ORIGIN
 1.4%; Score 37; DB 13; Length 658;
 Best Local Similarity 100.0%; Pred No. 8.4e-08;
 Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2600 TTTCATGCAATTATGATGTTTACATGCTGGAA 2636
 |||||
 Db 600 TTTCATGCAATTATGATGTTTACATGCTGGAA 564
 |||||
 RESULT 52
 BB612039 660 bp mRNA linear EST 26-OCT-2001
 LOCUS BB612039 660 bp mRNA linear EST 26-OCT-2001

DEFINITION BB612039 RIKEN full-length enriched, 15 days embryo head Mus musculus cDNA clone 4022435H11 5', mRNA sequence.
 ACCESSION BB612039
 VERSION BB612039.1 GI:16453118
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE 1 (bases 1 to 660)
 AUTHORS Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Horii,F., Ishii,Y., Ito,M., Kawai,J., Kono,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ono,H., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeeda,Y., Tanaka,T., Toyota,T., Muramatsu,M., and Hayashizaki,Y.
 TITLE RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
 JOURNAL Unpublished
 COMMENT Contact: Yoshitake Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suenho-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-re@gsc.riken.go.jp/
 URL: http://genome.gsc.riken.go.jp/
 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi,K., Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watanabe,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A., and Hayashizaki,Y.
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Kono,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,Y., and Hayashizaki,Y.
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo,S., Shinagawa,A., Saito,T., Kiyosawa,H., Yamanka,I., Atzawa,K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K., and Hayashizaki,Y.
 Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (http://genome.gsc.riken.go.jp) for further details.
 e mouse tissues.
 Location/Qualifiers
 1..660
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="4022435H11"
 /sex="mixed"
 /tissue_type="head"
 /dev_stage="15 days embryo"
 /lab_host="DHI0B"
 /clone_lib="RIKEN full-length enriched, 15 days embryo head"
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'.

GAGAGAGAGAGATCCAGAGCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATCTCGAGTTATTAATTAATCCCCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I"

BASE COUNT 185 a 141 c 160 g 174 t

Query Match 1.4%; Score 37; DB 10; Length 660;
Best Local Similarity 100.0%; Pred. No. 8.4e-08;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 306 ACAAGAAATGTCGATCCAGATGTCATCTTT 342
DB 163 ACAAGAAATGTCGATCCAGATGTCATCTTT 199

RESULT 53

CA327319 699 bp mRNA linear EST 27-NOV-2002

LOCUS UI-M-FY0-cy-p-01-0-UI.r1 NIH_BMAP_FY0 Mus musculus cDNA clone

IMAGE: 6826322 5', mRNA sequence.

CA327319

CA327319.1 GI:24545417

EST.

Mus musculus (house mouse)

SOURCE

ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 699)

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Dr. Jim Lin, University of Iowa

cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

This clone was contributed by the Brain Molecular Anatomy Project

(BMAP)

Seq primer: pyx-5.

Location/Qualifiers

1. .699

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE: 6826322"

/issue_type="whole brain"

/dev_stage="embryo 13.5, 14.5, 16.5, 17.5dpc"

/lab_host="DH10B (T1 phage resistant)"

/clone_id="NIH_BMAP_FY0"

/note="Organ: Brain; Vector: pyx-Asc; Site 1: Ecor I; Site 2: Not I; The library was constructed according

Bonafido, Lennon and Soares, Genome Research, 6:791-806,

1996. Denatured RNA was size fractionated on a 1% agarose

gel. First strand cDNA synthesis was primed with oligo-dT

primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated

with Ecor I adaptor, digested with NotI and then cloned

directionally into pyx-Asc vector. The library tag

sequence located between the Not I site and the polyA tail

is AGGAGAGCAG. This library was created for the University

Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the

Developing Mouse Nervous System', supported by National

BASE COUNT 197 a 165 c 156 g 180 t 1 others
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

Query Match 1.3%; Score 35; DB 14; Length 699;
Best Local Similarity 100.0%; Pred. No. 1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 986 GCTAGATAGTCCAGATATGCAAGATGACTTGG 1020
DB 382 GCTAGATAGTCCAGATATGCAAGATGACTTGG 416

RESULT 54 773 bp mRNA linear EST 12-SEP-2002

LOCUS BU515916

AGENCOUNT 10114062 NIH_MGC_134 Mus musculus cDNA clone

IMAGE: 6512603 5', mRNA sequence.

BU515916

BU515916.1 GI:22823442

EST.

Mus musculus (house mouse)

SOURCE

ORGANISM

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 773)

NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished

Contact: Robert Strausberg, Ph.D.

Email: cgabbs-remail.nih.gov

Tissue Procurement: Dr. David Rowe

cDNA Library Preparation: Invitrogen Corp

DNA Sequencing by: Agencourt Bioscience Corporation (LLNL)

Clone Distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLNL4084 row: P column: 12

High quality sequence stop: 616.

Location/Qualifiers

1. .773

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="IMAGE: 6512603"

/issue_type="undifferentiated limb"

/lab_host="DH10B (phage-resistant)"

/note="Vector: PCWV-SPORT6.1.cdb; Site 1: EcorV; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average

insert size 1.7 kb. Constructed by Reagen, Invitrogen

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

Corp. Note: this is a NIH_MGC Library."

KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Mammalia; Metazoa; Chordata; Vertebrata; Euteleostomi; Eukaryota; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 775)
NIR-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/BLINL at: <http://image.lnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project (BMAP)

FEATURES
Seq primer: PYX-5.
Location/Qualifiers
1..775
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE: 6405252"
/tissue_type="whole brain"
/dev_stage="embryo 12.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_1ib="NIR BMAP F00"
/note="Organ: Brain; Vector: PYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into PYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is TGAAGAGGCC. This library was created for the University Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT 221 a 184 c 165 g 200 t 5 others

ORIGIN

Query Match 1.3%; Score 35; DB 13; Length 775;
Best Local Similarity 100.0%; Pred. No. 1.1e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 986 GCTAGATAGTCAATATGCAAGATGACTTTGG 1020
|||||
Db 279 GCTAGATAGTCAATATGCAAGATGACTTTGG 313
|||||

RESULT 56
AA861394 374 bp mRNA linear EST 04-JUN-1999
LOCUS ak37c09.s1 Soares_testis_NHT Homo sapiens cDNA clone IMAGE:1408144
DEFINITION 3', mRNA sequence.
ACCESSION AA861394
VERSION AA861394.1 GI:2953534
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 374)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/BLINL at: www.bio.lnl.gov/bbrp/image/image.html
Insert Length: 1657 Std Error: 0.00
Seq primer: -40m3 fwd. ET from Amersham
High quality sequence stop: 364.
Location/Qualifiers
1..374
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1408144"
/sex="male"
/lab_host="DH10B"
/clone_1ib="Soares_testis_NHT"
/note="Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was prepared from mRNA obtained from Clontech Laboratories Inc., and primed with a Not I - oligo(dT) primer [5'-TGTTCACATCTGAAAGTGGAGCGCCGCCAATTTTCTTTTCTTTT 3'-]. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization to Cot5 and was constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 122 a 68 c 61 g 123 t

ORIGIN

Query Match 1.3%; Score 34; DB 9; Length 374;
Best Local Similarity 100.0%; Pred. No. 3e-06;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2676 GTAATTCTCAATAGGCTTTTCTTCTCAATACC 2709
|||||
Db 374 GTAATTCTCAATAGGCTTTTCTTCTCAATACC 341
|||||

RESULT 57
B1989239 600 bp mRNA linear EST 20-DEC-2001
LOCUS B1989239
DEFINITION 4032-91 Mouse E14.5 retina lambda ZAP II library Mus musculus cDNA, mRNA sequence.
ACCESSION B1989239
VERSION B1989239.1 GI:17960234
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 600)
Mu, X., Zhao, S., Pershad, R., Haieh, T., F., Scarpa, A., Wang, S.W., White, R.A., Beremand, P.D., Thomas, T.L., Gan, L. and Klein, W.H. Gene expression in the developing mouse retina by EST sequencing and microarray analysis
Nucleic Acids Res. 29 (24), 4983-4993 (2001)

JOURNAL MEDLINE 21671825
PUBMED 11812828
COMMENT Contact: Klein WH
Department of Biochemistry and Molecular Biology
University of Texas M.D. Anderson Cancer Center
Box 117, 1515 Holcombe Blvd., Houston, TX 77030, USA
Tel: 713 792 3646
Fax: 713 790 0329.
Location/Qualifiers

source 1..600
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="neural retina"
/dev_stage="embryonic day 14.5 post-fertilization"
/clone_lib="Mouse E14.5 retina lambda ZAP II library"

BASE COUNT 197 a 134 c 142 g 127 t
ORIGIN

Query Match 1.1%; Score 30; DB 12; Length 600;
Best Local Similarity 100.0%; Pred. No. 0.0005;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2600 TTTCATGCATTATGTCATGTTTACATG 2629
Db 392 TTTCATGCATTATGTCATGTTTACATG 421
|||||

RESULT 58
LOCUS BB628362 662 bp mRNA linear EST 31-AUG-2001
DEFINITION BB628362 RIKEN full-length enriched, adult male urinary bladder Mus musculus cDNA clone 9530096M16 5', mRNA sequence.
ACCESSION BB628362
VERSION BB628362.1 GI:15399354
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 662)
Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hitamoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyé, T., Muramatsu, M., and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
Unpublished
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.c.riken.go.jp, URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M., Yoneda, Y., Ishikawa, T., Osawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y., and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Yamanaka, I., Kiyosawa, H., Kondo, S., Saito, T., Shingawa, A., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., Arakawa, T., Ishii, Y., and Hayashizaki, Y.
Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct. Func. Genomics 2 pre, 172-186 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for

FEATURES
source
1..662
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone_lib="9530096M16"
/sex="male"
/tissue_type="urinary bladder"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, adult male urinary bladder"

BASE COUNT 221 a 135 c 153 g 152 t 1 others
ORIGIN

Query Match 1.1%; Score 30; DB 10; Length 662;
Best Local Similarity 100.0%; Pred. No. 0.00051;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2600 TTTCATGCATTATGTCATGTTTACATG 2629
Db 270 TTTCATGCATTATGTCATGTTTACATG 299
|||||

RESULT 59
LOCUS W98303 690 bp mRNA linear EST 16-JUL-1996
DEFINITION W98303 mg13h01.r1 Soares mouse embryo NBHE13.5 14.5 Mus musculus cDNA clone IMAGE:423697 5' similar to gb:X85993 M.musculus mRNA for semaphorin D (MOUSE);, mRNA sequence.
ACCESSION W98303
VERSION W98303.1 GI:1428279
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 690)
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theisinger, B., Wylie, T., Lennon, G., Soares, B., Wilson, R., and Waterston, R.
The WashU-HMI Mouse EST Project
Unpublished
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.

further details.

e mouse tissues.

Location/Qualifiers

Db 245 CATTGGAGAGACTTCTCAATAGATGA 273

RESULT 63

CD349263

DEFINITION

CD349263 729 bp mRNA linear EST 29-MAY-2003
UI-M-FY0-cfg-a-18-0-UI.r1 NIH BMAP_FY0 Mus musculus CDNA clone
IMAGE: 6854395 5', mRNA sequence.

ACCESSION

CD349263

VERSION

CD349263.1 GI:31140778

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE

NIH-MGC <http://mgc.nci.nih.gov/>.

AUTHORS

1 (bases 1 to 729)

TITLE

Unpublished

JOURNAL

Contact: Robert Strausberg, Ph.D.

COMMENT

Email: cgabs@mail.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
CDNA library preparation: Dr. M. Bento Soares, University of Iowa
CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
<http://genome.uiowa.edu/distribution/mouseefl.html>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

seg primer: pyx-5.

FEATURES

Location/Qualifiers

1..729

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6"

/db_xref="taxon:10090"

/clone="IMAGE: 6854395"

/tissue_type="whole brain"

/dev_stage="embryo 13.5, 14.5, 16.5, 17.5, 18.5, 19.5, 20.5, 21.5, 22.5, 23.5, 24.5, 25.5, 26.5, 27.5, 28.5, 29.5, 30.5, 31.5, 32.5, 33.5, 34.5, 35.5, 36.5, 37.5, 38.5, 39.5, 40.5, 41.5, 42.5, 43.5, 44.5, 45.5, 46.5, 47.5, 48.5, 49.5, 50.5, 51.5, 52.5, 53.5, 54.5, 55.5, 56.5, 57.5, 58.5, 59.5, 60.5, 61.5, 62.5, 63.5, 64.5, 65.5, 66.5, 67.5, 68.5, 69.5, 70.5, 71.5, 72.5, 73.5, 74.5, 75.5, 76.5, 77.5, 78.5, 79.5, 80.5, 81.5, 82.5, 83.5, 84.5, 85.5, 86.5, 87.5, 88.5, 89.5, 90.5, 91.5, 92.5, 93.5, 94.5, 95.5, 96.5, 97.5, 98.5, 99.5, 100.5, 101.5, 102.5, 103.5, 104.5, 105.5, 106.5, 107.5, 108.5, 109.5, 110.5, 111.5, 112.5, 113.5, 114.5, 115.5, 116.5, 117.5, 118.5, 119.5, 120.5, 121.5, 122.5, 123.5, 124.5, 125.5, 126.5, 127.5, 128.5, 129.5, 130.5, 131.5, 132.5, 133.5, 134.5, 135.5, 136.5, 137.5, 138.5, 139.5, 140.5, 141.5, 142.5, 143.5, 144.5, 145.5, 146.5, 147.5, 148.5, 149.5, 150.5, 151.5, 152.5, 153.5, 154.5, 155.5, 156.5, 157.5, 158.5, 159.5, 160.5, 161.5, 162.5, 163.5, 164.5, 165.5, 166.5, 167.5, 168.5, 169.5, 170.5, 171.5, 172.5, 173.5, 174.5, 175.5, 176.5, 177.5, 178.5, 179.5, 180.5, 181.5, 182.5, 183.5, 184.5, 185.5, 186.5, 187.5, 188.5, 189.5, 190.5, 191.5, 192.5, 193.5, 194.5, 195.5, 196.5, 197.5, 198.5, 199.5, 200.5, 201.5, 202.5, 203.5, 204.5, 205.5, 206.5, 207.5, 208.5, 209.5, 210.5, 211.5, 212.5, 213.5, 214.5, 215.5, 216.5, 217.5, 218.5, 219.5, 220.5, 221.5, 222.5, 223.5, 224.5, 225.5, 226.5, 227.5, 228.5, 229.5, 230.5, 231.5, 232.5, 233.5, 234.5, 235.5, 236.5, 237.5, 238.5, 239.5, 240.5, 241.5, 242.5, 243.5, 244.5, 245.5, 246.5, 247.5, 248.5, 249.5, 250.5, 251.5, 252.5, 253.5, 254.5, 255.5, 256.5, 257.5, 258.5, 259.5, 260.5, 261.5, 262.5, 263.5, 264.5, 265.5, 266.5, 267.5, 268.5, 269.5, 270.5, 271.5, 272.5, 273.5, 274.5, 275.5, 276.5, 277.5, 278.5, 279.5, 280.5, 281.5, 282.5, 283.5, 284.5, 285.5, 286.5, 287.5, 288.5, 289.5, 290.5, 291.5, 292.5, 293.5, 294.5, 295.5, 296.5, 297.5, 298.5, 299.5, 300.5, 301.5, 302.5, 303.5, 304.5, 305.5, 306.5, 307.5, 308.5, 309.5, 310.5, 311.5, 312.5, 313.5, 314.5, 315.5, 316.5, 317.5, 318.5, 319.5, 320.5, 321.5, 322.5, 323.5, 324.5, 325.5, 326.5, 327.5, 328.5, 329.5, 330.5, 331.5, 332.5, 333.5, 334.5, 335.5, 336.5, 337.5, 338.5, 339.5, 340.5, 341.5, 342.5, 343.5, 344.5, 345.5, 346.5, 347.5, 348.5, 349.5, 350.5, 351.5, 352.5, 353.5, 354.5, 355.5, 356.5, 357.5, 358.5, 359.5, 360.5, 361.5, 362.5, 363.5, 364.5, 365.5, 366.5, 367.5, 368.5, 369.5, 370.5, 371.5, 372.5, 373.5, 374.5, 375.5, 376.5, 377.5, 378.5, 379.5, 380.5, 381.5, 382.5, 383.5, 384.5, 385.5, 386.5, 387.5, 388.5, 389.5, 390.5, 391.5, 392.5, 393.5, 394.5, 395.5, 396.5, 397.5, 398.5, 399.5, 400.5, 401.5, 402.5, 403.5, 404.5, 405.5, 406.5, 407.5, 408.5, 409.5, 410.5, 411.5, 412.5, 413.5, 414.5, 415.5, 416.5, 417.5, 418.5, 419.5, 420.5, 421.5, 422.5, 423.5, 424.5, 425.5, 426.5, 427.5, 428.5, 429.5, 430.5, 431.5, 432.5, 433.5, 434.5, 435.5, 436.5, 437.5, 438.5, 439.5, 440.5, 441.5, 442.5, 443.5, 444.5, 445.5, 446.5, 447.5, 448.5, 449.5, 450.5, 451.5, 452.5, 453.5, 454.5, 455.5, 456.5, 457.5, 458.5, 459.5, 460.5, 461.5, 462.5, 463.5, 464.5, 465.5, 466.5, 467.5, 468.5, 469.5, 470.5, 471.5, 472.5, 473.5, 474.5, 475.5, 476.5, 477.5, 478.5, 479.5, 480.5, 481.5, 482.5, 483.5, 484.5, 485.5, 486.5, 487.5, 488.5, 489.5, 490.5, 491.5, 492.5, 493.5, 494.5, 495.5, 496.5, 497.5, 498.5, 499.5, 500.5, 501.5, 502.5, 503.5, 504.5, 505.5, 506.5, 507.5, 508.5, 509.5, 510.5, 511.5, 512.5, 513.5, 514.5, 515.5, 516.5, 517.5, 518.5, 519.5, 520.5, 521.5, 522.5, 523.5, 524.5, 525.5, 526.5, 527.5, 528.5, 529.5, 530.5, 531.5, 532.5, 533.5, 534.5, 535.5, 536.5, 537.5, 538.5, 539.5, 540.5, 541.5, 542.5, 543.5, 544.5, 545.5, 546.5, 547.5, 548.5, 549.5, 550.5, 551.5, 552.5, 553.5, 554.5, 555.5, 556.5, 557.5, 558.5, 559.5, 560.5, 561.5, 562.5, 563.5, 564.5, 565.5, 566.5, 567.5, 568.5, 569.5, 570.5, 571.5, 572.5, 573.5, 574.5, 575.5, 576.5, 577.5, 578.5, 579.5, 580.5, 581.5, 582.5, 583.5, 584.5, 585.5, 586.5, 587.5, 588.5, 589.5, 590.5, 591.5, 592.5, 593.5, 594.5, 595.5, 596.5, 597.5, 598.5, 599.5, 600.5, 601.5, 602.5, 603.5, 604.5, 605.5, 606.5, 607.5, 608.5, 609.5, 610.5, 611.5, 612.5, 613.5, 614.5, 615.5, 616.5, 617.5, 618.5, 619.5, 620.5, 621.5, 622.5, 623.5, 624.5, 625.5, 626.5, 627.5, 628.5, 629.5, 630.5, 631.5, 632.5, 633.5, 634.5, 635.5, 636.5, 637.5, 638.5, 639.5, 640.5, 641.5, 642.5, 643.5, 644.5, 645.5, 646.5, 647.5, 648.5, 649.5, 650.5, 651.5, 652.5, 653.5, 654.5, 655.5, 656.5, 657.5, 658.5, 659.5, 660.5, 661.5, 662.5, 663.5, 664.5, 665.5, 666.5, 667.5, 668.5, 669.5, 670.5, 671.5, 672.5, 673.5, 674.5, 675.5, 676.5, 677.5, 678.5, 679.5, 680.5, 681.5, 682.5, 683.5, 684.5, 685.5, 686.5, 687.5, 688.5, 689.5, 690.5, 691.5, 692.5, 693.5, 694.5, 695.5, 696.5, 697.5, 698.5, 699.5, 700.5, 701.5, 702.5, 703.5, 704.5, 705.5, 706.5, 707.5, 708.5, 709.5, 710.5, 711.5, 712.5, 713.5, 714.5, 715.5, 716.5, 717.5, 718.5, 719.5, 720.5, 721.5, 722.5, 723.5, 724.5, 725.5, 726.5, 727.5, 728.5, 729.5

/lab_host="DH10B (T1 phage resistant)"

/note="Organ: Brain; Vector: pyx-Asc; Site: 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction. ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pyx-Asc vector. The library tag
is AGCGAGACG. This library was created for the University
Iowa Brain Anatomy Project (BMAP). 'Gene Discovery in the
Developing Mouse Nervous System', supported by National
Institute of Mental Health (NIMH), Hemn Chin, Ph.D.,
program coordinator."

/clone_lib="NIH_BMAP_FY0"

/db_xref="taxon:10090"

/tissue_type="whole brain"

/dev_stage="embryo 13.5, 14.5, 16.5, 17.5, 18.5, 19.5, 20.5, 21.5, 22.5, 23.5, 24.5, 25.5, 26.5, 27.5, 28.5, 29.5, 30.5, 31.5, 32.5, 33.5, 34.5, 35.5, 36.5, 37.5, 38.5, 39.5, 40.5, 41.5, 42.5, 43.5, 44.5, 45.5, 46.5, 47.5, 48.5, 49.5, 50.5, 51.5, 52.5, 53.5, 54.5, 55.5, 56.5, 57.5, 58.5, 59.5, 60.5, 61.5, 62.5, 63.5, 64.5, 65.5, 66.5, 67.5, 68.5, 69.5, 70.5, 71.5, 72.5, 73.5, 74.5, 75.5, 76.5, 77.5, 78.5, 79.5, 80.5, 81.5, 82.5, 83.5, 84.5, 85.5, 86.5, 87.5, 88.5, 89.5, 90.5, 91.5, 92.5, 93.5, 94.5, 95.5, 96.5, 97.5, 98.5, 99.5, 100.5, 101.5, 102.5, 103.5, 104.5, 105.5, 106.5, 107.5, 108.5, 109.5, 110.5, 111.5, 112.5, 113.5, 114.5, 115.5, 116.5, 117.5, 118.5, 119.5, 120.5, 121.5, 122.5, 123.5, 124.5, 125.5, 126.5, 127.5, 128.5, 129.5, 130.5, 131.5, 132.5, 133.5, 134.5, 135.5, 136.5, 137.5, 138.5, 139.5, 140.5, 141.5, 142.5, 143.5, 144.5, 145.5, 146.5, 147.5, 148.5, 149.5, 150.5, 151.5, 152.5, 153.5, 154.5, 155.5, 156.5, 157.5, 158.5, 159.5, 160.5, 161.5, 162.5, 163.5, 164.5, 165.5, 166.5, 167.5, 168.5, 169.5, 170.5, 171.5, 172.5, 173.5, 174.5, 175.5, 176.5, 177.5, 178.5, 179.5, 180.5, 181.5, 182.5, 183.5, 184.5, 185.5, 186.5, 187.5, 188.5, 189.5, 190.5, 191.5, 192.5, 193.5, 194.5, 195.5, 196.5, 197.5, 198.5, 199.5, 200.5, 201.5, 202.5, 203.5, 204.5, 205.5, 206.5, 207.5, 208.5, 209.5, 210.5, 211.5, 212.5, 213.5, 214.5, 215.5, 216.5, 217.5, 218.5, 219.5, 220.5, 221.5, 222.5, 223.5, 224.5, 225.5, 226.5, 227.5, 228.5, 229.5, 230.5, 231.5, 232.5, 233.5, 234.5, 235.5, 236.5, 237.5, 238.5, 239.5, 240.5, 241.5, 242.5, 243.5, 244.5, 245.5, 246.5, 247.5, 248.5, 249.5, 250.5, 251.5, 252.5, 253.5, 254.5, 255.5, 256.5, 257.5, 258.5, 259.5, 260.5, 261.5, 262.5, 263.5, 264.5, 265.5, 266.5, 267.5, 268.5, 269.5, 270.5, 271.5, 272.5, 273.5, 274.5, 275.5, 276.5, 277.5, 278.5, 279.5, 280.5, 281.5, 282.5, 283.5, 284.5, 285.5, 286.5, 287.5, 288.5, 289.5, 290.5, 291.5, 292.5, 293.5, 294.5, 295.5, 296.5, 297.5, 298.5, 299.5, 300.5, 301.5, 302.5, 303.5, 304.5, 305.5, 306.5, 307.5, 308.5, 309.5, 310.5, 311.5, 312.5, 313.5, 314.5, 315.5, 316.5, 317.5, 318.5, 319.5, 320.5, 321.5, 322.5, 323.5, 324.5, 325.5, 326.5, 327.5, 328.5, 329.5, 330.5, 331.5, 332.5, 333.5, 334.5, 335.5, 336.5, 337.5, 338.5, 339.5, 340.5, 341.5, 342.5, 343.5, 344.5, 345.5, 346.5, 347.5, 348.5, 349.5, 350.5, 351.5, 352.5, 353.5, 354.5, 355.5, 356.5, 357.5, 358.5, 359.5, 360.5, 361.5, 362.5, 363.5, 364.5, 365.5, 366.5, 367.5, 368.5, 369.5, 370.5, 371.5, 372.5, 373.5, 374.5, 375.5, 376.5, 377.5, 378.5, 379.5, 380.5, 381.5, 382.5, 383.5, 384.5, 385.5, 386.5, 387.5, 388.5, 389.5, 390.5, 391.5, 392.5, 393.5, 394.5, 395.5, 396.5, 397.5, 398.5, 399.5, 400.5, 401.5, 402.5, 403.5, 404.5, 405.5, 406.5, 407.5, 408.5, 409.5, 410.5, 411.5, 412.5, 413.5, 414.5, 415.5, 416.5, 417.5, 418.5, 419.5, 420.5, 421.5, 422.5, 423.5, 424.5, 425.5, 426.5, 427.5, 428.5, 429.5, 430.5, 431.5, 432.5, 433.5, 434.5, 435.5, 436.5, 437.5, 438.5, 439.5, 440.5, 441.5, 442.5, 443.5, 444.5, 445.5, 446.5, 447.5, 448.5, 449.5, 450.5, 451.5, 452.5, 453.5, 454.5, 455.5, 456.5, 457.5, 458.5, 459.5, 460.5, 461.5, 462.5, 463.5, 464.5, 465.5, 466.5, 467.5, 468.5, 469.5, 470.5, 471.5, 472.5, 473.5, 474.5, 475.5, 476.5, 477.5, 478.5, 479.5, 480.5, 481.5, 482.5, 483.5, 484.5, 485.5, 486.5, 487.5, 488.5, 489.5, 490.5, 491.5, 492.5, 493.5, 494.5, 495.5, 496.5, 497.5, 498.5, 499.5, 500.5, 501.5, 502.5, 503.5, 504.5, 505.5, 506.5, 507.5, 508.5, 509.5, 510.5, 511.5, 512.5, 513.5, 514.5, 515.5, 516.5, 517.5, 518.5, 519.5, 520.5, 521.5, 522.5, 523.5, 524.5, 525.5, 526.5, 527.5, 528.5, 529.5, 530.5, 531.5, 532.5, 533.5, 534.5, 535.5, 536.5, 537.5, 538.5, 539.5, 540.5, 541.5, 542.5, 543.5, 544.5, 545.5, 546.5, 547.5, 548.5, 549.5, 550.5, 551.5, 552.5, 553.5, 554.5, 555.5, 556.5, 557.5, 558.5, 559.5, 560.5, 561.5, 562.5, 563.5, 564.5, 565.5, 566.5, 567.5, 568.5, 569.5, 570.5, 571.5, 572.5, 573.5, 574.5, 575.5, 576.5, 577.5, 578.5, 579.5, 580.5, 581.5, 582.5, 583.5, 584.5, 585.5, 586.5, 587.5, 588.5, 589.5, 590.5, 591.5, 592.5, 593.5, 594.5, 595.5, 596.5, 597.5, 598.5, 599.5, 600.5, 601.5, 602.5, 603.5, 604.5, 605.5, 606.5, 607.5, 608.5, 609.5, 610.5, 611.5, 612.5, 613.5, 614.5, 615.5, 616.5, 617.5, 618.5, 619.5, 620.5, 621.5, 622.5, 623.5, 624.5, 625.5, 626.5, 627.5, 628.5, 629.5, 630.5, 631.5, 632.5, 633.5, 634.5, 635.5, 636.5, 637.5, 638.5, 639.5, 640.5, 641.5, 642.5, 643.5, 644.5, 645.5, 646.5, 647.5, 648.5, 649.5, 650.5, 651.5, 652.5, 653.5, 654.5, 655.5, 656.5, 657.5, 658.5, 659.5, 660.5, 661.5, 662.5, 663.5, 664.5, 665.5, 666.5, 667.5, 668.5, 669.5, 670.5, 671.5, 672.5, 673.5, 674.5, 675.5, 676.5, 677.5, 678.5, 679.5, 680.5, 681.5, 682.5, 683.5, 684.5, 685.5, 686.5, 687.5, 688.5, 689.5, 690.5, 691.5, 692.5, 693.5, 694.5, 695.5, 696.5, 697.5, 698.5, 699.5, 700.5, 701.5, 702.5, 703.5, 704.5, 705.5, 706.5, 707.5, 708.5, 709.5, 710.5, 711.5, 712.5, 713.5, 714.5, 715.5, 716.5, 717.5, 718.5, 719.5, 720.5, 721.5, 722.5, 723.5, 724.5, 725.5, 726.5, 727.5, 728.5, 729.5

/lab_host="DH10B (T1 phage resistant)"

/note="Organ: Brain; Vector: pyx-Asc; Site: 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1

ORIGIN

Query Match 1.0%; Score 27; DB 10; Length 446;
 Best Local Similarity 100.0%; Pred. No. 0.019;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 TTTCAGATTGTTGAACCTCTCTGGC 144
 |||||
 377 TTTCAGATTGTTGAACCTCTCTGGC 403

RESULT 65
 AL918018 539 bp mRNA linear EST 18-SEP-2002
 LOCUS AL918018 PUR-Z1+Z2 Danio rerio CDNA clone 042-A05-2, mRNA sequence.
 DEFINITION AL918018
 ACCESSION AL918018
 VERSION AL918018.1 GI:23183316
 KEYWORDS EST.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio

REFERENCE
 AUTHORS Lee,S., Ruan,H., Ma,W., Xu,M., Lo,J., He,Y., Liu,F., Eun,A., Wen,Z.
 and Peng,J.
 15,000 unique zebrafish EST clusters from two cDNA libraries
 and Peng,J.
 Contact: Peng J
 Lab of Functional Genomics
 Institute of Molecular and Cell Biology
 30 Medical Drive, Singapore, 117609, Singapore
 Email: pengj@imcb.a-star.edu.sg
 Clone requests: pengj@imcb.a-star.edu.sg
 Location/Qualifiers

FEATURES
 source 1..539
 /organism="Danio rerio"
 /mol_type="mRNA"
 /strain="local wildtype"
 /db_xref="taxon:7955"
 /clone="042-A05-2"
 /tissue_type="whole embryo or fish"
 /dev_stage="mixed stages"
 /clone_idb="PUR-Z1+Z2"
 /clone_1lb="PUR-Z1+Z2"

BASE COUNT 141 a 126 c 139 g 133 t
 ORIGIN

Query Match 1.0%; Score 27; DB 9; Length 539;
 Best Local Similarity 100.0%; Pred. No. 0.02;
 Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 500 GATGAATGCAGTGGCTGGAAAGAC 526
 |||||
 250 GATGAATGCAGTGGCTGGAAAGAC 276

RESULT 66
 BG926854 685 bp mRNA linear EST 06-NOV-2001
 LOCUS BG926854
 DEFINITION HNC20-1-B9.R HNC (Human Normal Cartilage) Homo sapiens cDNA, mRNA
 sequence.
 ACCESSION BG926854
 VERSION BG926854.1 GI:14321377
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Kumar,S., Connor,J.R., Dodds,R.A., Halsey,W., Van Horn,M., Mao,J.,
 Sathe,G., Mui,P., Agarwal,P., Badger,A.M., Lee,D.C., Gowen,M. and
 Laird,M.W.
 Identification and initial characterization of 5000 expressed

sequenced tags (ESTs) each from adult human normal and
 osteoarthritic cartilage cDNA libraries
 Osteoarthr. Cartil. 9 (7), 641-653 (2001)
 MEDLINE 21482651
 PUBMED 11597177
 COMMENT Contact: Sanjay Kumar
 UW2109
 GlaxoSmithKline
 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
 Tel: 610-270-7245
 Fax: 610-270-5598
 Email: sanjay.kumar-1@gsk.com
 Seq primer: T7.
 Location/Qualifiers

FEATURES
 source 1..685
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="cartilage"
 /lab_host="E.coli DH10 B"
 /clone_1lb="HNC (Human Normal Cartilage)"
 /note="Vector: pSPORT I; Site_1: SalI; Site_2: NotI;
 Directional"

BASE COUNT 178 a 150 c 193 g 164 t
 ORIGIN

Query Match 0.9%; Score 25; DB 12; Length 685;
 Best Local Similarity 100.0%; Pred. No. 0.26;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGGCTCCGGAGTAGG 81
 |||||
 77 GTGACCCACGGCTCCGGAGTAGG 101

RESULT 67
 CD203422 556 bp mRNA linear EST 21-MAY-2003
 LOCUS CD203422
 DEFINITION Ls AM1_08D06 T7 Litomosoides sigmodontis adult male 1 Litomosoides
 sigmodontis cDNA clone Ls AM1_08D06 5' similar to P90699 (P90699)
 Putative RNA binding protein, mRNA sequence.
 CD203422
 VERSION CD203422.1 GI:30931437
 KEYWORDS EST.
 SOURCE Litomosoides sigmodontis
 ORGANISM Litomosoides sigmodontis

REFERENCE
 AUTHORS LeGeoff,J., Thomas,F., Allen,J. and Blaxter,M.L.
 1 (bases 1 to 556)
 Genes expressed in adult male Litomosoides sigmodontis
 Onchocercidae; Litomosoides.
 Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;

COMMENT
 Title: Blaxter ML
 Contact: Blaxter ML
 Institute of Cell, Animal and Population Biology
 University of Edinburgh
 Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
 3JT, UK.
 Tel: +44 131 650 6760
 Fax: +44 131 670 5450
 Email: mark.blaxter@ed.ac.uk
 The library was constructed by laetitia LeGeoff in Judith Allen's
 laboratory at ICAPB, University of Edinburgh. Sequencing performed
 by Fran Thomas, in Mark Blaxter's laboratory, ICAPB, University of
 Edinburgh. The sequence contained a Polya tail (trimmed)

PCR PRIMERS
 FORWARD: M13L
 BACKWARD: M13R
 Place: 08 row: D column: 06
 Seq primer: T7
 High quality sequence start: 8
 High quality sequence stop: 533.
 Location/Qualifiers

FEATURES
 source 1..556

```

/organism="Litomosoides sigmodontis"
/mol_type="mRNA"
/db_xref="taxon:42156"
/clone="Ls_AM1_08006"
/sex="male"
/dev_stage="adult"
/clone_lib="Litomosoides sigmodontis adult male 1"
/notes="Vector: pSPORT-1; Site 1: Sal I; Site 2: Not I; The library was constructed by Laetitia Legeroff in Judith Allen's laboratory at ICAPB, University of Edinburgh. The library is directional, with the 5' end of inserts adjacent to the Sal I site; sequencing with the T7 primer thus gives 5' ESTs."
BASE COUNT      112 a      114 c      148 g      182 t
ORIGIN

Query Match      0.9%; Score 24; DB 14; Length 556;
Best Local Similarity 100.0%; Pred. No. 0.85;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      57 GTGACCCACGCGTCCGGAGTAG 80
      |||
Db      12 GTGACCCACGCGTCCGGAGTAG 35

RESULT 68      598 bp      mRNA      linear      EST 21-MAY-2003
CD203073      Ls_AM1_02E03_T7 Litomosoides sigmodontis adult male 1 Litomosoides
LOCUS      sigmodontis cDNA clone Ls_AM1_02E03 5' similar to MSP2_ONCVO
ACCESSION      (P13263) Major sperm protein 2 (MSP2), mRNA sequence.
VERSION      CD203073
KEYWORDS      EST.
SOURCE      Litomosoides sigmodontis
ORGANISM      Litomosoides sigmodontis
      Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
      Onchocercidae; Litomosoides.
REFERENCE      1 (bases 1 to 598)
AUTHORS      Legeroff, L., Thomas, F., Allen, J. and Blaxter, M.L.
TITLE      Genes expressed in adult male Litomosoides sigmodontis
JOURNAL      Unpublished
COMMENT      Contact: Blaxter ML
      Institute of Cell, Animal and Population Biology
      University of Edinburgh
      Ashworth Labs, King's Buildings, West Mains Road, Edinburgh, EH9
      3JT, UK.
      Tel: +44 131 650 6760
      Fax: +44 131 670 5450
      Email: mark.blaxter@ed.ac.uk
      The library was constructed by Laetitia Legeroff in Judith Allen's
      laboratory at ICAPB, University of Edinburgh. Sequencing performed
      by Fran Thomas, in Mark Blaxter's laboratory, ICAPB, University of
      Edinburgh. The sequence contained a PolyA tail (trimmed)
PCR PRIMERS
FORWARD: M13L
BACKWARD: M13R
Plate: 02 row: E column: 03
Seq primer: T7
High quality sequence start: 12
High quality sequence stop: 531.
Location/Qualifiers
1..598
/organism="Litomosoides sigmodontis"
/mol_type="mRNA"
/db_xref="taxon:42156"
/clone="Ls_AM1_02E03"
/sex="male"
/dev_stage="adult"
/clone_lib="Litomosoides sigmodontis adult male 1"
/notes="Vector: pSPORT-1; Site 1: Sal I; Site 2: Not I; The library was constructed by Laetitia Legeroff in Judith Allen's laboratory at ICAPB, University of Edinburgh. The

```

```

library is directional, with the 5' end of inserts
adjacent to the Sal I site; sequencing with the T7 primer
thus gives 5' ESTs."
BASE COUNT      156 a      151 c      157 g      134 t
ORIGIN

Query Match      0.9%; Score 24; DB 14; Length 598;
Best Local Similarity 100.0%; Pred. No. 0.87;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      57 GTGACCCACGCGTCCGGAGTAG 80
      |||
Db      32 GTGACCCACGCGTCCGGAGTAG 55

RESULT 69      614 bp      mRNA      linear      EST 14-NOV-2002
CA502146      WHE4043_C11_P21ZT wheat meiotic anther cDNA library Triticum
LOCUS      aestivum cDNA clone WHE4043_C11_P21, mRNA sequence.
DEFINITION      CA502146
ACCESSION      CA502146
VERSION      CA502146.1 GI:24993106
KEYWORDS      EST.
SOURCE      Triticum aestivum (bread wheat)
ORGANISM      Triticum aestivum
      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
      Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
      ; Triticeae; Triticum.
REFERENCE      1 (bases 1 to 614)
AUTHORS      Anderson, O.D., Chao, S., Crossman, C., Langridge, P., Lazo, G.R., Pham
      J., Rausch, C.J., Sutton, T., Woo, J. and Wilson, C.
TITLE      The structure and function of the expressed portion of the wheat
      genomes - Meiotic anther cDNA library
JOURNAL      Unpublished
COMMENT      Contact: Olin Anderson
      US Department of Agriculture, Agriculture Research Service, Pacific
      West Area, Western Regional Research Center
      800 Buchanan Street, Albany, CA 94710, USA
      Tel: 5105595773
      Fax: 5105595818
      Email: oanderson@pw.usda.gov
      Sequences have been trimmed to remove vector sequence and low
      quality sequence with phred score less than 20
      Seq primer: T7 primer.
Location/Qualifiers
1..614
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE4043_C11_P21"
/tissue_type="Anther"
/dev_stage="Meiotic stages pre-meiosis-metaphase I"
/lab_host="E. coli DH10B"
/clone_lib="Wheat meiotic anther cDNA library"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; Plants
were grown in a glasshouse. Anther meiotic stage was
determined by removing anthers from individual primary
florets. One anther was sacrificed for microscopic staging
and if determined to be between (and including) meiotic
stages pre-meiosis and metaphase I, the remaining two
anthers were collected and pooled for library
construction. The tissue, total RNA, and poly(A) RNA were
prepared, cDNA synthesised, and directionally ligated into
pSPORT1 by Tim Sutton in the P Langridge Lab at the
Department of Plant Science, University of Adelaide, Waite
Campus, Australia. Average insert size 1.5kb. Plasmid DNA
preparations and DNA sequencing were performed in the OD
Anderson lab (all other authors)."
BASE COUNT      168 a      164 c      163 g      119 t
ORIGIN

Query Match      0.9%; Score 24; DB 14; Length 614;

```


Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source

```
1. 851
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hixex"
/db_xref="taxon:9031"
/clone="CHST3399e23"
/dev_stage="36"
/lab_host="DH10B"
/clone_1ib="CSECHN63"
/note="Organ: heads; Vector: pBluescript II KS(+); Site_1:
EcoRI; Site_2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunted, ligated to NotI adapters, digested with EcoRI
, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."
```

BASE COUNT
ORIGIN

295 a 153 c 154 g 249 t

Query Match 0.9%; Score 24; DB 13; Length 851;
Best Local Similarity 100.0%; Pred. No. 0.96;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2578 AATGCAATATACATGACCTTTT 2601

128 AATGCAATATACATGACCTTTT 151

RESULT 73

BU267223

LOCUS 897 bp mRNA linear EST 26-NOV-2002
DEFINITION 603815106r1 CSECHN52 Gallus gallus cDNA clone CHEST80711 5', mRNA

ACCESSION BU267223

VERSION BU267223.1 GI:25538173

KEYWORDS EST.

SOURCE Gallus gallus (chicken)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Archosauria; Aves; Neognathae; Galliformes; Phasianidae;

AUTHORS 1 (bases 1 to 897)

TITLE Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,

JOURNAL Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.

MEDLINE A Comprehensive Collection of Chicken cDNAs

PUBMED Curr. Biol. 12 (22), 1965-1969 (2002)

COMMENT Contact: Simon Hubbard

Department of Biomolecular Sciences

University of Manchester Institute of Science and Technology (UMIST)

PO Box 88, Manchester, M60 1QD, UK

Tel: 01612008930

Fax: 01612360409

Email: Simon.Hubbard@umist.ac.uk.

Location/Qualifiers

1. 897

/organism="Gallus gallus"

/mol_type="mRNA"

/strain="White Leghorn, Hixex"

/db_xref="taxon:9031"

/clone="CHST80711"

/dev_stage="22"

/lab_host="DH10B"

/clone_1ib="CSECHN52"

/note="Organ: limbs; Vector: pBluescript II KS(+); Site_1:

EcoRI; Site_2: NotI; This normalized library was

constructed from 1 million independent clones. cDNA

synthesis was initiated using an oligo(dT) primer, using

methylated C in the first strand synthesis reaction.

Following this first strand reaction, double-stranded cDNA

was blunted, ligated to NotI adapters, digested with EcoRI

, size-selected, and cloned into the NotI and EcoRI

compatible sites of a custom modified MCS of the

pBluescript (KS+) vector. The library was normalized in 2

rounds using conditions adapted from Soares et al., PNAS

(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6

(1996): 791, except that a significantly longer

reannealing hybridization was used."

BASE COUNT 306 a 172 c 174 g 245 t

ORIGIN

Query Match 0.9%; Score 24; DB 13; Length 897;

Best Local Similarity 100.0%; Pred. No. 0.97;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2578 AATGCAATATACATGACCTTTT 2601

192 AATGCAATATACATGACCTTTT 215

RESULT 74

AM080828

LOCUS 184 bp mRNA linear EST 14-OCT-1999

DEFINITION xc38a05.x1 NCI CGAP Co20 Homo sapiens cDNA clone IMAGE:2586512 3'

ACCESSION AM080828

VERSION AM080828.1 GI:6035980

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE Mammalia; Eutheria; Primates; Carnivora; Hominoidea; Homo.

1 (bases 1 to 184)

TITLE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Unpublished

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgap@nci.nih.gov

Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael

R. Emmert-Buck, M.D., Ph.D., cDNA Library Preparation: Life

Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The

I.M.A.G.E. Consortium DNA Sequencing by: Washington University

Genome Sequencing Center

Clone Distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/ILNI at:

www.bio.lnlnl.gov/bbrp/image/image.html

Seq primer: -40UP from Gibco.

Location/Qualifiers

1. 184

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:2586512"

/issue_type="moderately differentiated adenocarcinoma"

/lab_host="DH10B"

/clone_1ib="NCI CGAP Co20"

/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: SalI;

Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Normalized to Cot 500. Average insert size 1.1kb.
 Normalized version of NCI_CGAP_CO18. Library constructed
 by Life Technologies."

BASE COUNT 50 a 46 c 38 g 50 t

Query Match 0.8%; Score 23; DB 9; Length 184;
 Best Local Similarity 100.0%; Pred. No. 2.2;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAGTA 79
 ||||||||||||||||||
 DB 179 GTGACCCACGCGTCCGGAGTA 157

RESULT 75
 BG898214 401 bp mRNA linear EST 06-NOV-2001
 LOCUS HOA38-1-B10 HOA (Human Osteoarthritic Cartilage) Homo sapiens cDNA,
 DEFINITION mRNA sequence.

ACCESSION BG898214
 VERSION BG898214.1 GI:14308463
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 401)
 AUTHORS Kumar,S., Connor,J.R., Dodds,R.A., Halsey,W., Van Horn,M., Mao,J.,
 Sathie,G., Nui,P., Agarwal,P., Badger,A.M., Lee,J.C., Gowen,M. and
 Lark,M.W.

TITLE Identification and initial characterization of 5000 expressed
 sequenced tags (ESTs) each from adult human normal and
 osteoarthritic cartilage cDNA libraries
 JOURNAL Osteoarthr. Cartil. 9 (7), 641-653 (2001)
 MEDLINE 21482651
 PUBMED 11597177

COMMENT Contact: Sanjay Kumar
 UW2109

GlaxoSmithKline
 709 Swedeland Road, P.O. Box 1539, King of Prussia, PA 19406, USA
 Tel: 610-270-7245
 Fax: 610-270-5598
 Email: sanjay.kumar-1@sk.com
 Seq primer: T7.

FEATURES
 Location/Qualifiers

1..401
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue_type="cartilage"
 /lab_host="E. coli DH10 B"
 /clone_lib="HOA (Human Osteoarthritic Cartilage)"
 /note="Vector: pSPORT 1; Site_1: SalI; Site_2: NotI;
 Directional"

BASE COUNT 67 a 132 c 130 g 72 t

ORIGIN

Query Match 0.8%; Score 23; DB 12; Length 401;
 Best Local Similarity 100.0%; Pred. No. 2.7;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAGTA 79
 ||||||||||||||||||
 DB 69 GTGACCCACGCGTCCGGAGTA 91

Search completed: July 31, 2003, 23:52:32
 Job time : 5037 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 17:17:31 ; Search time 679 Seconds
(without alignments)
10769.923 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709

Sequence: 1 aaattctattatcatgcatg.....agcttttttctcattacc 2709

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : N Geneseq.19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2709	100.0	2709	22	AAH47049
2	2508	92.6	2530	19	AAST7567
3	2508	92.6	2530	24	AAST7562
4	2508	92.6	2601	16	AAO87442
5	1253	46.3	1481	16	AAO92331
6	405	15.0	456	23	ABS45213
7	405	15.0	456	24	ABS19795
8	303	11.2	354	22	AAI29182

9	303	11.2	354	25	ABZ33368	Human colon tumour
10	292	10.8	496	23	ABZ32213	Human liver single
11	292	10.8	496	24	ABZ32213	Human genome-deriv
12	160	5.9	172	22	ABA71205	Human foetal liver
13	160	5.9	172	22	AAK19504	Human brain expres
14	160	5.9	172	22	AAK45495	Human bone marrow
15	160	5.9	172	22	AAI51439	Probe #420125 used
16	160	5.9	172	23	ABZ45419	Human liver single
17	160	5.9	172	24	ABZ19763	Human genome-deriv
18	160	5.9	484	22	ABA58707	Human foetal liver
19	160	5.9	484	22	AAK06842	Human brain expres
20	160	5.9	484	22	AAK32561	Human bone marrow
21	160	5.9	484	22	AAI18386	Probe #7072 used t
22	160	5.9	484	23	ABZ32272	Human liver single
23	160	5.9	484	23	ABZ07351	Human genome-deriv
24	144	5.3	446	22	ABZ07722	Human brain expres
25	144	5.3	446	23	ABZ33341	Human liver single
26	144	5.3	446	23	ABZ08426	Human genome-deriv
27	123	4.5	123	22	AAK20366	Human brain expres
28	123	4.5	123	23	ABZ46246	Human liver single
29	123	4.5	123	24	ABZ20840	Human genome-deriv
30	122	4.5	456	22	ABA58557	Human foetal liver
31	122	4.5	456	22	ABA27590	Probe #6056 for ge
32	122	4.5	456	22	AAK0676	Human brain expres
33	122	4.5	456	22	AAK32372	Human bone marrow
34	122	4.5	456	22	AAI38218	Probe #5904 used t
35	122	4.5	456	23	ABZ32080	Human liver single
36	122	4.5	456	24	ABZ07155	Human genome-deriv
37	101	3.7	101	22	ABA71291	Human foetal liver
38	101	3.7	101	22	ABA7573	Probe #16039 for g
39	101	3.7	101	22	AAK1593	Human brain expres
40	101	3.7	101	22	AAK45596	Human bone marrow
41	101	3.7	101	22	AAI51533	Probe #40219 used
42	101	3.7	101	23	ABZ45286	Human liver single
43	101	3.7	101	24	ABZ19868	Human genome-deriv
44	71	2.6	2136	22	AAH47053	Human maspin CDNA
45	60	2.2	229	21	AAK02557	Human secreted pro
46	60	2.2	229	24	ABN43098	Human leukocyte ge
47	50	1.8	50	24	ABZ04740	Human lactalbumin
48	31	1.1	4661	24	AAZ28273	Alpha-lactalbumin
49	31	1.1	4661	24	AAZ28312	Alpha-lactalbumin
50	29	1.1	29	25	ABZ25421	Human semaphorin-3
51	27	1.0	9729	22	AAK63093	Nucleotide sequenc
52	27	1.0	12473	22	AAK63092	Nucleotide sequenc
53	26	1.0	1313	22	ABZ52430	Human polynucleoti
54	25	0.9	25	25	ABZ25416	Human semaphorin-3
55	25	0.9	1399	22	AAH47052	Murine maspin CDNA
56	24	0.9	24	25	ABZ25415	Human semaphorin-3
57	24	0.9	8417	21	AAK66819	Reporter gene cons
58	24	0.9	8684	21	AAK66820	Reporter gene cons
59	23	0.8	29	25	ABZ25419	Human semaphorin-3
60	23	0.8	175	24	ABK87383	CDNA fragment for
61	23	0.8	240	24	ABK87385	CDNA fragment for
62	23	0.8	287	24	ABK87379	CDNA encoding part
63	23	0.8	326	22	AAI18251	Human breast canc
64	23	0.8	335	24	ABK87384	CDNA fragment for
65	23	0.8	339	24	ABK87380	CDNA encoding part
66	23	0.8	352	24	ABK87382	CDNA encoding part
67	23	0.8	527	24	ABK87381	CDNA encoding part
68	23	0.8	631	24	AAK63284	Human immune/haema
69	23	0.8	661	24	ABK87378	CDNA encoding/haema
70	22	0.8	115	22	AAH36207	Human colon cancer
71	22	0.8	370	22	AAK61215	Human immune/haema
72	22	0.8	471	22	AAI29660	Human polynucleoti
73	22	0.8	711	24	ABZ77002	Frog embryonic gen
74	22	0.8	748	24	ABZ77128	Frog embryonic gen
75	22	0.8	2112	21	AAH67682	Human secreted pro
76	22	0.8	2112	24	AAZ33705	Human secreted pro
77	22	0.8	5732	24	AAZ28269	Mouse mammary tumo
78	22	0.8	5732	24	AAV81412	Mouse mammary tumo
79	22	0.8	7311	20	AAV81412	Plasmid pLNMK used
80	22	0.8	7591	22	AAK63094	Nucleotide sequenc
81	22	0.8	7885	20	AAV81413	Plasmid pLNMK12x us

CC 82 22 0.8 9072 20 AAX07558 Synthetic modified
C 83 22 0.8 9183 24 AAD28270 Alpha-lactalbumin
C 84 22 0.8 9183 24 AAD28309 Alpha-lactalbumin
C 85 21 0.8 29 25 ABZ25420 Human semaphorin-3
C 86 21 0.8 332 25 ABX61182 Arabidopsis thaliana
C 87 21 0.8 341 22 AAL19287 Human breast cancer
C 88 21 0.8 385 22 ABA08285 Human protein-Tyr
C 89 21 0.8 442 21 AAF16393 Human prostate can
C 90 21 0.8 457 25 ABX39208 Bovine EST associa
C 91 21 0.8 489 22 ABK57109 Human immune/haema
C 92 21 0.8 578 25 ABZ53641 Aspergillus oryzae
C 93 21 0.8 717 22 AHA34766 Human colon cancer
C 94 21 0.8 730 24 ABS77215 Frog embryonic gen
C 95 21 0.8 745 24 ABS77448 Frog embryonic gen
C 96 21 0.8 749 24 ABS77126 Frog embryonic gen
C 97 21 0.8 749 24 ABS77195 Frog embryonic gen
C 98 21 0.8 750 24 ABS77132 Frog embryonic gen
C 99 21 0.8 750 24 ABS77271 Frog embryonic gen
C 100 21 0.8 781 24 ABS77150 Frog embryonic gen

ALIGNMENTS

RESULT 1
ID AHA47049 standard; DNA; 2709 BP.
XX AHA47049;
AC AHA47049;
DT 29-OCT-2001 (first entry)
DE Semaphorin D cDNA sequence.
XX
XX Drug resistance; resistance gene; semaphorin D; B94; mel-14 antigen;
XX 24p3; proliferin; maspin; cancer; cytostatic; gene therapy; ss.
XX Homo sapiens.
XX WO200155455-A2.
XX 02-AUG-2001.
XX 31-JAN-2001; 2001WO-US03161.
XX 31-JAN-2000; 2000US-0179191.
XX (MILL-) MILLENNIUM PHARM INC.
XX (JINS/) JIN S.
XX Jin S;
XX WPI: 2001-488799/53.
XX
XX Determining if a compound modulates the drug resistance of a cell,
XX comprises determining the expression or activity level of a resistance
XX sequence in a cell in the presence of the test compound -
XX
XX Example 1; Fig 1A-B; 79pp; English.
XX
XX The invention relates to a method of determining whether a test compound
XX modulates the drug resistance of a cell that comprises determining the
XX expression or activity level of resistance genes (e.g. semaphorin D, B94,
XX mel-14 antigen, 24p3, proliferin or maspin) in a cell in the presence of
XX the test compound, and comparing its expression or activity level in a
XX cell without the test compound. The drug resistant sequences are useful
XX in identifying drug resistant cells, in screening methods directed to the
XX identification of compounds that can modulate the drug resistance of a
XX cell type or multiple cell types. An isolated resistance protein can be
XX used as an immunogen to generate antibodies that bind the resistance
XX protein. Resistance nucleic acids may be inserted into vectors and used
XX as gene therapy vectors. An anti-resistance protein antibody may be used
XX to isolate a resistance protein, or facilitate the purification of

CC natural resistance protein from cells and of recombinantly produced
CC resistance protein expressed in host cells. The methods are useful for
CC treating a subject having a disorder, such as a drug-resistance cancer,
CC characterized by aberrant resistance sequence expression or activity by
CC administering to the subject a resistance modulator. The present sequence
CC represents a semaphorin cDNA sequence, whose expression was increased in
CC drug resistant EMT6 tumours.
XX
SQ Sequence 2709 BP; 831 A; 560 C; 620 G; 698 T; 0 other;
Query Match 100.0%; Score 2709; DB 22; Length 2709;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AATCTTTATTTATCGATGTTAACAACCTTAGTAATGATGCCACCTCGAGGGTTCG 60
DB 1 AATCTTTATTTATCGATGTTAACAACCTTAGTAATGATGCCACCTCGAGGGTTCG 60
QY 61 ACCCAGCGGTCGGAGTAGGTGAGCTCGGCTGTTCCCATTTGTCAGCCAGTCTATT 120
DB 61 ACCCAGCGGTCGGAGTAGGTGAGCTCGGCTGTTCCCATTTGTCAGCCAGTCTATT 120
QY 121 CCAGATTGTTGAACCTTCTGCGCCGCAATATACAGAGAGACTTAAGCCAGCAAGG 180
DB 121 CCAGATTGTTGAACCTTCTGCGCCGCAATATACAGAGAGACTTAAGCCAGCAAGG 180
QY 181 GACCTACAGCGCTCGACGACATGAGCTGTTAATGATGATGATGATGATGATGAT 240
DB 181 GACCTACAGCGCTCGACGACATGAGCTGTTAATGATGATGATGATGATGATGAT 240
QY 241 ATTACTTACAGCAAGCAACATATCAGATGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 241 ATTACTTACAGCAAGCAACATATCAGATGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 301 ATCTTACAAAGAAATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 360
DB 301 ATCTTACAAAGAAATGTTGATGATGATGATGATGATGATGATGATGATGATGAT 360
QY 361 CTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 361 CTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
QY 421 GGATCACAATTTTTCATTCGACCTGTTAATATCAGAGATTTTCAAAAGATGTTG 480
DB 421 GGATCACAATTTTTCATTCGACCTGTTAATATCAGAGATTTTCAAAAGATGTTG 480
QY 481 AGTATCTTACCGCAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 481 AGTATCTTACCGCAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 540
QY 541 TGCTAATTTTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
DB 541 TGCTAATTTTCAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
QY 601 GGGGGCTTTTTCATTCATTTGCACTTATGAAATTTGACATCATCTGAGGACAAT 660
DB 601 GGGGGCTTTTTCATTCATTTGCACTTATGAAATTTGACATCATCTGAGGACAAT 660
QY 661 TTTTAACTGGAGAACTCAATTTTGAACCGGCTGGAGAGTCCATATGACCTTAA 720
DB 661 TTTTAACTGGAGAACTCAATTTTGAACCGGCTGGAGAGTCCATATGACCTTAA 720
QY 721 GCTGCTACAGCATCCCTTTTAAATGAGAGATTAATCTGGAATCTGAGCTGATTT 780
DB 721 GCTGCTACAGCATCCCTTTTAAATGAGAGATTAATCTGGAATCTGAGCTGATTT 780
QY 781 TATGGGCGAGACTTTTCTATCTTCCGAACTTTGGGACACACCAATCAGACAGA 840
DB 781 TATGGGCGAGACTTTTCTATCTTCCGAACTTTGGGACACACCAATCAGACAGA 840
QY 841 GCAGCATGATTCAGAGTGGCTCAATGATCCAAAGTTATTAATGATGATGATGAT 900
DB 841 GCAGCATGATTCAGAGTGGCTCAATGATCCAAAGTTATTAATGATGATGATGAT 900

OY	901	GAGTGCATTCCTGGAAGTAGCAAAAGTAATCTTTTCTCCGGAAAAATGCCAATACATGCG	960
Db	901	GAGTGCAATCTCGAAGAAGCAAAAGTAATCTTTTCTCCGGAAAAATCCAATAAGATGG	960
OY	961	AGAACACTCTGGAAAAGCTACTCACCCTGAAATAGGTGAGATATGCAAGATGACTTTGG	102
Db	961	AGAACACTCTGGAAAAGCTACTCACCCTGAAATAGGTGAGATATGCAAGATGACTTTGG	102
OY	1021	AGGGCACAGAAATCTGTGTGATTAATTGACACAATTCTTCAAAGCTGTCTGATTTGGCTC	108
Db	1021	AGGGCACAGAAATCTGTGTGATTAATTGACACAATTCTTCAAAGCTGTCTGATTTGGCTC	108
OY	1081	AGTGCAGAGTCCAAATNGCATTTGACACATCTATTTTGTGAACTGACAGATGTATTCCTAAT	114
Db	1081	AGTGCAGAGTCCAAATNGCATTTGACACATCTATTTTGTGAACTGACAGATGTATTCCTAAT	114
OY	1141	GAACTTTAAAGATCCTTAAAAATCCAGTTGTATATGGAAGTGTYYACACTTCCAGTAACAT	120
Db	1141	GAACTTTAAAGATCCTTAAAAATCCAGTTGTATATGGAAGTGTYYACACTTCCAGTAACAT	120
OY	1201	TTTTCAAGGGATAGCGGTGTATGTATATGATGATGATATGTGAAAGGATGTTCTTGG	126
Db	1201	TTTTCAAGGGATAGCGGTGTATGTATATGATGATGATATGTGAAAGGATGTTCTTGG	126
OY	1261	TCCATATGCCCAACAGGSATGAGCCCACTATCAATGGAGCTTATCAAGSAAAGATCCC	132
Db	1261	TCCATATGCCCAACAGGSATGAGCCCACTATCAATGGAGCTTATCAAGSAAAGATCCC	132
OY	1321	CTATCCAGGGCCAGGAATTGTGCCAGCAAAACATTTGGTGTTTGACTCTACAAAGAA	138
Db	1321	CTATCCAGGGCCAGGAATTGTGCCAGCAAAACATTTGGTGTTTGACTCTACAAAGAA	138
OY	1381	CCTTCCTGATGATGTATTAATCCTTTGCAAGAAAGTATCCAGGCATGTACAATCCAGTGT	144
Db	1381	CCTTCCTGATGATGTATTAATCCTTTGCAAGAAAGTATCCAGGCATGTACAATCCAGTGT	144
OY	1441	TCCATGAAACAATCGCCCCATATGATCAAAAACGATGTAAATTATCAATTTTACACAAT	150
Db	1441	TCCATGAAACAATCGCCCCATATGATCAAAAACGATGTAAATTATCAATTTTACACAAT	150
OY	1501	TGTGCTAACCAGAGTGCATGCAAGAGTAGACAGATATGTTATGTTTATGCGAACAG	156
Db	1501	TGTGCTAACCAGAGTGCATGCAAGAGTAGACAGATATGTTATGTTTATGCGAACAG	156
OY	1561	TGTTGGGACCGTTCTTAAAGTAGTTCATCTCTAAGAGAGACTGGTATGATTTTGAAGA	162
Db	1561	TGTTGGGACCGTTCTTAAAGTAGTTCATCTCTAAGAGAGACTGGTATGATTTTGAAGA	162
OY	1621	GGTTCTGCTGGAAGAAATGACAGTTTTCGGGAAACCGACTGCTATTTGACAAATGAGCT	168
Db	1621	GGTTCTGCTGGAAGAAATGACAGTTTTCGGGAAACCGACTGCTATTTGACAAATGAGCT	168
OY	1681	TTCCACTAAGACAGCAACAATAATATGSGTTCAACGGCTGGGGTGGTCCCACTCCCTTT	174
Db	1681	TTCCACTAAGACAGCAACAATAATATGSGTTCAACGGCTGGGGTGGTCCCACTCCCTTT	174
OY	1741	AACCCGATGTATATTTCACGGGAAAGCGTGTCTGAGTGTTCCTCGCCGACCCGACCTTA	180
Db	1741	AACCCGATGTATATTTCACGGGAAAGCGTGTCTGAGTGTTCCTCGCCGACCCGACCTTA	180
OY	1801	CTGTGCTTGGGATGTTCTGCACTGTTCTCGCTATTTTCCCACTGCAAAAGAGACGACAAAG	186
Db	1801	CTGTGCTTGGGATGTTCTGCACTGTTCTCGCTATTTTCCCACTGCAAAAGAGACGACAAAG	186
OY	1861	ACGACAAATATTAAGAAATGAGACCCACTGACTCACTGTCAGACTTACACATGATTA	192
Db	1861	ACGACAAATATTAAGAAATGAGACCCACTGACTCACTGTCAGACTTACACATGATTA	192
OY	1921	TCACCATGGCCACAGCCCTGAAAGAGAAATCATATGTTGTAGAGAAATAGTAGCAATT	198
Db	1921	TCACCATGGCCACAGCCCTGAAAGAGAAATCATATGTTGTAGAGAAATAGTAGCAATT	198

Oy		1981	TTTGGAAATGCACTCCGAAGTGCGACAGAGCCCTGGTCATTATGGCAATTCCAGAGCCGAA	2040
Db		1981	TTTTGAATGCAGTCCGAAGTGCACAGAGCCCTGGTCATTATGGCAATTCCAGAGCCGAA	2040
Oy		2041	TGAAGAGCGAAAAAGAAGATCAGAGTGGATGATCATTCATCCAGAGACAGATCAAAGGCT	2100
Db		2041	TGAAGAGCGAAAAAGAAGATCAGAGTGGATGATCATTCATCCAGAGACAGATCAAAGGCT	2100
Oy		2101	TCTGTACGTAGTCTTACACAAGAGATTCAAGGCATTAACCTCTGCCATGCGGTGGACA	2160
Db		2101	TCTGTACGTAGTCTTACACAAGAGATTCAAGGCATTAACCTCTGCCATGCGGTGGACA	2160
Oy		2161	TGGGTTCATCAAACTCTTCTTAAAGTACCTTGGAAAGTCAATGACACAGAGCATTTGGA	2220
Db		2161	TGGGTTCATCAAACTCTTCTTAAAGTACCTTGGAAAGTCAATGACACAGAGCATTTGGA	2220
Oy		2221	AGAACTTCTTCATTAAGAATGATGATGATGAGATGGCTTAAGCCCAAGAAANTGTCATATG	2280
Db		2221	AGAACTTCTTCATTAAGAATGATGATGATGAGATGGCTTAAGCCCAAGAAANTGTCATATG	2280
Oy		2281	CATGACACTTAGCCAGAGAGGTCTGGTACAGACCTTCATGACGCTCATCAACCACCCA	2340
Db		2281	CATGACACTTAGCCAGAGAGGTCTGGTACAGACCTTCATGACGCTCATCAACCACCCA	2340
Oy		2341	TCTCAACACGATGATGATGATCTGTGAAACAAGTTGAAAAGGACCGAAGAACAGCTCG	2400
Db		2341	TCTCAACACGATGATGATGATCTGTGAAACAAGTTGAAAAGGACCGAAGAACAGCTCG	2400
Oy		2401	GCAAAAGCCAGACATPACCCTCCAGGGAAACAGTAAACAATNGAAGACCTTACAAGAAATPA	2460
Db		2401	GCAAAAGCCAGACATPACCCTCCAGGGAAACAGTAAACAATNGAAGACCTTACAAGAAATPA	2460
Oy		2461	GAAAGGTGAAAACAGAGAGACCCACGAATTTGAGAGGGACCCAGAGTGTCTGAGCTGC	2520
Db		2461	GAAAGGTGAAAACAGAGAGACCCACGAATTTGAGAGGGACCCAGAGTGTCTGAGCTGC	2520
Oy		2521	ATTACTCTTAGAAACTTCAACAAGTGAATACTTGCTTAGACATTAACCTGAAAAACAA	2580
Db		2521	ATTACTCTTAGAAACTTCAACAAGTGAATACTTGCTTAGACATTAACCTGAAAAACAA	2580
Oy		2581	TGCAATATPACATGAACTTTTTCATGGCATTAATGATGATGTTTCAATGGTGGAAATTG	2640
Db		2581	TGCAATATPACATGAACTTTTTCATGGCATTAATGATGATGTTTCAATGGTGGAAATTG	2640
Oy		2641	AGCTGAGTCCACCAATTAATTAATCATGAGTAACTTTCCATATAGGCTTTT	2700
Db		2641	AGCTGAGTCCACCAATTAATTAATCATGAGTAACTTTCCATATAGGCTTTT	2700
Oy		2701	CCTAATACC 2709	
Db		2701	CCTAATACC 2709	
RESULT 2				
ID	AAK75767		standard; DNA; 2530 BP.	
XX	AAK75767;			
XX				
DT	22-JUL-1999	(first entry)		
XX				
DE	Human semaphorin III DNA.			
XX				
KM	Human; beta-amyloid precursor protein; beta-APP; diagnosis; cancer;			
KM	familial Alzheimer's disease; age-related disease; neurodegenerative disorder;			
KM	Alzheimer's disease; Down's syndrome; myotonic dystrophy; neuronal;			
KM	Huntington's disease; multiple sclerosis; alcoholic liver disease;			
KM	diabetes mellitus type II; microtubule associated protein; Tau; Big Tau;			
KM	ubiquitin B; apolipoprotein E; MAP2; neurofilament-L; neurofilament-M;			
KM	neurofilament-F; presenilin I; presenilin II; cellular tumour antigen;			
KM	glial fibrillary acidic protein; GRAP; p53; semaphorin III; HUPF-1;			
KM	bcl-2; B-cell leukemia/lymphoma 2 proto-oncogene; NMEP-C; NSP-A;			
KM	high mobility group protein-C; neuroendocrine specific protein A; ss.			

1573 TCTTAAAGTAGTTTCAATCTCTAAGAGACCTTGATGATTTAGAGAGGCTTCTGCTGA 1632
 1389 TCTTAAAGTAGTTTCAATCTCTAAGAGACCTTGATGATTTAGAGAGGCTTCTGCTGA 1448
 1633 AGAATGACAGTTTGGGGAACCGAGCTGATTTTCAAGATGAGGCTTCCACATGCA 1692
 1449 AGAATGACAGTTTGGGGAACCGAGCTGATTTTCAAGATGAGGCTTCCACATGCA 1508
 1693 GCAACACTATATATATGTTTCAACCGAGCTGAGGCTTCCAGCTCCCTTTACACCGGTGGA 1752
 1509 GCAACACTATATATATGTTTCAACCGAGCTGAGGCTTCCAGCTCCCTTTACACCGGTGGA 1566
 1753 TATTTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGACCTTTACTGTGTTGGGA 1812
 1569 TATTTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGACCTTTACTGTGTTGGGA 1628
 1813 TGTGTTCTGATGTTTCTGCTATTTTCCACCTGCAAGAGACCGACACAAACGACATAT 1872
 1629 TGTGTTCTGATGTTTCTGCTATTTTCCACCTGCAAGAGACCGACACAAACGACATAT 1688
 1873 AAGAAATGAGAGACCGACTGACTCAGTGTTCAGACTTACACCATGATATACCATGGCCA 1932
 1689 AAGAAATGAGAGACCGACTGACTCAGTGTTCAGACTTACACCATGATATACCATGGCCA 1748
 1933 CAGCCCTGAAGAGAGATCATCTATGTTGATAGAGATGATGACATTTTGGAAATGCGAG 1992
 1749 CAGCCCTGAAGAGAGATCATCTATGTTGATAGAGATGATGACATTTTGGAAATGCGAG 1808
 1993 TCCGAAGTCCGAGAGAGGCTGTGTTATTTGGCAATTCAGAGGCGCAATTTGAAGCGGAA 2052
 1809 TCCGAAGTCCGAGAGAGGCTGTGTTATTTGGCAATTCAGAGGCGCAATTTGAAGCGGAA 1868
 2053 AGAAGAGATCAGAGTGTGATCATATCATCAGAGACAGTCAAGGCTTCTGCTACGTAG 2112
 1869 AGAAGAGATCAGAGTGTGATCATATCATCAGAGACAGTCAAGGCTTCTGCTACGTAG 1928
 2113 TCTACACAGAGAGATTCAGGCAATTTACCTCTGCGAGTGGTGAACATGAGGTTATCA 2172
 1929 TCTACACAGAGAGATTCAGGCAATTTACCTCTGCGAGTGGTGAACATGAGGTTATCA 1988
 2173 AACTCTTTTAAAGTAAACCTCTGAGAGTATGACACAGACATTTGGAAGAACTTTCTCA 2232
 1989 AACTCTTTTAAAGTAAACCTCTGAGAGTATGACACAGACATTTGGAAGAACTTTCTCA 2048
 2233 TAAAGATGATGATGAGAGTGGCTCTAAGACCAAGAAATGTCATATAGATGACACTTAC 2292
 2049 TAAAGATGATGATGAGAGTGGCTCTAAGACCAAGAAATGTCATATAGATGACACTTAC 2108
 2293 CCAGAGAGTGTGATCAGAGACTTCATGACGCTCATCAACCAACCCCAATCTCAACAGCAT 2352
 2109 CCAGAGAGTGTGATCAGAGACTTCATGACGCTCATCAACCAACCCCAATCTCAACAGCAT 2168
 2353 GAGTAGAGTGTGTAACAAGTTTGAAAAAGGACCGAAAAACAAGCTGCGCAAGGCCAGG 2412
 2169 GAGTAGAGTGTGTAACAAGTTTGAAAAAGGACCGAAAAACAAGCTGCGCAAGGCCAGG 2228
 2413 ACATACCCCGAGGAAACAGTAACTAATGAGACCTTAAAGAAATTAAGAAAGGTAGAAA 2472
 2229 ACATACCCCGAGGAAACAGTAACTAATGAGACCTTAAAGAAATTAAGAAAGGTAGAAA 2288
 2473 CAGAGAGACCAAGTATGAGAGGCAACCAAGAGTGTGAGGCTGATTAACCTCTAGA 2532
 2289 CAGAGAGACCAAGTATGAGAGGCAACCAAGAGTGTGAGGCTGATTAACCTCTAGA 2348
 2533 AACCTCAAAACAAGTAACTGCTTACACATTAATCTGAAAAACAATTCATATACAT 2592
 2349 AACCTCAAAACAAGTAACTGCTTACACATTAATCTGAAAAACAATTCATATACAT 2408
 2593 GAACTTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2652
 2409 GAACTTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2468
 2653 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2700

DB 2469 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2516
 RESULT 3
 ABS76512
 ID ABS76512 standard; cDNA; 2530 BP.
 AC ABS76512;
 XX 11-DEC-2002 (first entry)
 DE cDNA encoding human ovarian cancer marker M473.
 KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;
 KW central nervous system disorder; bacterial meningitis; viral meningitis;
 KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
 KW brain herniation; inflammation; encephalitis; testicular disorder;
 KW nontuberculous granulomatous orchitis; connective tissue disorder;
 KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
 KW histological type; carcinogenic; ovarian cancer marker; gene; ss.
 XX Homo sapiens.
 OS WO200271928-A2.
 PN 19-SEP-2002.
 PD 14-MAR-2002; 2002WO-US07826.
 PF 14-MAR-2001; 2001US-276025P.
 PR 14-MAR-2001; 2001US-276025P.
 PR 10-AUG-2001; 2001US-311732P.
 PR 19-SEP-2001; 2001US-323580P.
 PR 26-SEP-2001; 2001US-324967P.
 PR 26-SEP-2001; 2001US-325102P.
 PR 26-SEP-2001; 2001US-325149P.
 XX (MILL-) MILLENNIUM PHARM INC.
 PA Monahan JF, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG;
 PI Meyers RE, Morrissey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;
 PI Baet RC, Lu K, Schmandt RE, Zhao X, Glatt K;
 XX WPI: 2002-723277/78.
 DR P-PSDB; ABG95413.
 PT Assessing whether a patient is afflicted with ovarian cancer, useful in
 PT assessing the stage or progression of the disease, comprises comparing
 PT the expression level of a cancer marker in a sample from a patient and
 PT from a non cancer patient -
 XX Disclosure; Page 400; 481pp; English.
 PS The present invention relates to a new method for assessing whether a
 XX patient is afflicted with ovarian cancer. The method involves comparing
 CC the expression level of a marker in a patient sample and the normal level
 CC of expression of the marker in a control non-ovarian cancer sample, where
 CC the marker is selected from 363 cancer markers described in the
 CC specification. The method of the invention is useful in diagnosing or
 CC characterizing cancer, in detecting the presence of cancer as early as
 CC possible, and the recurrence of ovarian cancer. The method may also be of
 CC particular use with patients having an enhanced risk of developing
 CC ovarian cancer (e.g. patients having a familial history of ovarian
 CC cancer). The cancer markers may be used in the management and treatment
 CC of e.g. brain and central nervous system disorders (e.g. bacterial and
 CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain
 CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
 CC inflammations (e.g. bacterial or viral meningitis or encephalitis),
 CC testicular disorders (e.g. nontuberculous granulomatous orchitis),
 CC connective tissue disorders, or heart disorders (e.g. ischaemic heart
 CC disease or atherosclerosis). The compositions and methods may also be
 CC used in assessing the histological type of neoplasm associated with

CC ovarian cancer, monitoring the progression of ovarian cancer,
CC determining whether ovarian cancer has metastasized or is likely to
CC metastasize, selecting a composition for inhibiting ovarian cancer,
CC assessing the ovarian carcinogenic potential of a compound, or
CC inhibiting ovarian cancer or at risk of developing ovarian cancer. The
CC present nucleic acid sequence encodes one of the ovarian cancer markers
described in the invention.

XX
SQ Sequence 2530 BP; 786 A; 518 C; 576 G; 650 T; 0 other;

Query Match 92.6%; Score 2508; DB 24; Length 2530;

Best Local Similarity 100.0%; Pred.No.0;

Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

193 CTGACGACATGGGCTGGTAACTAGATTGTCTGCTTTCTGGGAGATTAATTACTTACAC 252
9 CTGACGACATGGGCTGGTAACTAGATTGTCTGCTTTCTGGGAGATTAATTACTTACAC 68
253 AAGAGCAACTATCAGAAATGGGAGAAACAATGTGCCAAGGCTGAAATTAATCTTACAAAGA 312
69 AAGAGCAACTATCAGAAATGGGAGAAACAATGTGCCAAGGCTGAAATTAATCTTACAAAGA 128
313 AATGTGGAAATCCAAATGTATCACTTTCAATGGCTGGCCAAAGCTCCAGTTATCA 372
129 AATGTGGAAATCCAAATGTATCACTTTCAATGGCTGGCCAAAGCTCCAGTTATCA 188
372 TACCTTCTTTGGATGAGAAACGAGTAGGCTGTATGTGGAGCAAGATCAATAT 432
189 TACCTTCTTTGGATGAGAAACGAGTAGGCTGTATGTGGAGCAAGATCAATAT 248
433 TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGATCTTTAC 492
249 TTCAATTCGACCTGGTTAATATCAAGGATTTTCAAAAGATTTGTGGCCAGATCTTTAC 308
493 CAGAAGATGTAATGGAATGGGCTGGAAAGACATCCCTGAAAGATGTGCTAATTCAT 552
309 CAGAAGATGTAATGGAATGGGCTGGAAAGACATCCCTGAAAGATGTGCTAATTCAT 368
553 CAGGACTTAAGCATATATACAGACTCACTTGTACGCTGTGAAACGGGGCTTTTCA 612
369 CAGGACTTAAGCATATATACAGACTCACTTGTACGCTGTGAAACGGGGCTTTTCA 428
613 TCCAAATTTGCACCTTACATTTGAATGGAATCTCTGAGAGCAATATTTTAAAGCTGA 672
429 TCCAAATTTGCACCTTACATTTGAATGGAATCTCTGAGAGCAATATTTTAAAGCTGA 488
673 GAACCTCACATTTTGAAGACGGCCGTGGAGAGAGTCCATATGACCCCTAACCTGTCACAC 732
489 GAACCTCACATTTTGAAGACGGCCGTGGAGAGAGTCCATATGACCCCTAACCTGTCACAC 548
733 ATCCCTTTAATAGATGAGAAATTAATCTGGAATCTGAGCTGATTTATGGGGGAGA 792
549 ATCCCTTTAATAGATGAGAAATTAATCTGGAATCTGAGCTGATTTATGGGGGAGA 608
793 CTTTGCTATCTTCCGAATCTTTGGGACACCAATCAGAGACAGAGCAGATGATTC 852
609 CTTTGCTATCTTCCGAATCTTTGGGACACCAATCAGAGACAGAGCAGATGATTC 668
853 CAGGTGGCTCATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGAATTC 912
669 CAGGTGGCTCATGATCCAAAGTTCAATAGTCCCACTCATCTCAGAGAGTGAATTC 728
913 TGAAGATGCAAGATATATCTTTCTTCCGTGAAATGCAATAGATGAGAGACATCTGG 972
729 TGAAGATGCAAGATATATCTTTCTTCCGTGAAATGCAATAGATGAGAGACATCTGG 788
973 AAAAGCTACACGCTAGATAGTTCAGATATGCAAGATGACTTTGGAGGGGACAGAG 1032
789 AAAAGCTACACGCTAGATAGTTCAGATATGCAAGATGACTTTGGAGGGGACAGAG 848
1033 TCTGTGTAATTAATGCAACATTTCTCAAAGCTGTCTGATTTGCTCAGTGCAGGTTC 1092
849 TCTGTGTAATTAATGCAACATTTCTCAAAGCTGTCTGATTTGCTCAGTGCAGGTTC 908

1093 AATGGCAATGCACTCACTTATGATGAACCTGAGAGATGATTTCTTAATGAATTTAAAGA 1152
909 AATGGCAATGCACTCACTTATGATGAACCTGAGAGATGATTTCTTAATGAATTTAAAGA 968
1153 TCTTAAAAATCCAGTTGATATGAGGTGTTTACAGCTTCCAGTAAATTTTCAAGGATC 1212
969 TCTTAAAAATCCAGTTGATATGAGGTGTTTACAGCTTCCAGTAAATTTTCAAGGATC 1028
1213 AGCCGTGTATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1272
1029 AGCCGTGTATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1088
1273 CAGGATGAGACCCCACTATCAATGGGCTCTTATCAAGAAAGTCCCTATCCAGGCT 1332
1089 CAGGATGAGACCCCACTATCAATGGGCTCTTATCAAGAAAGTCCCTATCCAGGCT 1148
1333 AGGAATCTTCCAGCAAAACATTTGGTGTGTTTACCTTACAAAGACCTTCTGATGA 1392
1149 AGGAATCTTCCAGCAAAACATTTGGTGTGTTTACCTTACAAAGACCTTCTGATGA 1208
1393 TGTATTAACCTTTGCAAGATCATCCAGCATGATCAATCAAGTTTCTATGAACA 1452
1209 TGTATTAACCTTTGCAAGATCATCCAGCATGATCAATCAAGTTTCTATGAACA 1268
1453 TCGCCCAATAGATCAAAACGAGATGTAATTAATCAATTTACAAATTTGCTAGACCG 1512
1269 TCGCCCAATAGATCAAAACGAGATGTAATTAATCAATTTACAAATTTGCTAGACCG 1328
1513 AGTGAATGCAAGATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1572
1329 AGTGAATGCAAGATGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1388
1573 TCTTAAGATGTTCAATTTCTTAAGAGAGACTTGTATGATTAAGAGAGTTCTGCTGA 1632
1389 TCTTAAGATGTTCAATTTCTTAAGAGAGACTTGTATGATTAAGAGAGTTCTGCTGA 1448
1633 AGAAATGCAAGTTTTCGGGAAACGAGCTGTAATTTACAAATGAGACTTTTCACTAACCA 1692
1449 AGAAATGCAAGTTTTCGGGAAACGAGCTGTAATTTACAAATGAGACTTTTCACTAACCA 1508
1693 GCAACACTATATATGTTTCAACGCTGGGGTGGCCAGCTCCTTTACACCGGTGGA 1752
1509 GCAACACTATATATGTTTCAACGCTGGGGTGGCCAGCTCCTTTACACCGGTGGA 1568
1753 TATTTACGGGAAGGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1812
1569 TATTTACGGGAAGGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1628
1813 TGTTCCTGATGTTCTGCTATATTTTCCATGCAAGAGACGCAACAGAGCAAGATAT 1872
1629 TGTTCCTGATGTTCTGCTATATTTTCCATGCAAGAGAGCGCAACAGAGATAT 1688
1873 AAGAAATGAGACCCCACTGATCTCACTGTTCAAGCTTACATCAATGATATCACATGGCA 1932
1689 AAGAAATGAGACCCCACTGATCTCACTGTTCAAGCTTACATCAATGATATCACATGGCA 1748
1933 CAGCCCTGAAGAGAAATCATCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1992
1749 CAGCCCTGAAGAGAAATCATCTATGATGATGATGATGATGATGATGATGATGATGATGAT 1808
1993 TCCGAATGCGAGAGAGGCTGTCTATTTGGCAATTCAGAGGCGCAATGAGAGCGAAA 2052
1809 TCCGAATGCGAGAGAGGCTGTCTATTTGGCAATTCAGAGGCGCAATGAGAGCGAAA 1868
2053 AGAAGATGCAAGTGTATGATATCATATCATAGAGACAGATCAAGGCTTCTGCTAGAG 2112
1869 AGAAGATGCAAGTGTATGATATCATATCATAGAGACAGATCAAGGCTTCTGCTAGAG 1928
2113 TCTTAAAGAGAAATTAAGGCAATTAATCTGCAAGGCTGAGAAATGAGGTTCAATCA 2172
1929 TCTTAAAGAGAAATTAAGGCAATTAATCTGCAAGGCTGAGAAATGAGGTTCAATCA 1988

QY 2173 AACTCTTCTTAAGTAACTCCCTGGAAAGTCAATTCAGACAGAGCAATTTGGAAGAATCTTCTCA 2232
 DB 1989 AACTCTTCTTAAGTAACTCCCTGGAAAGTCAATTCAGACAGAGCAATTTGGAAGAATCTTCTCA 2048
 QY 2223 TAAAGATGATGAGATGCTCTTAAGACCAAGAAATGTCATATGACATGACACCTAG 2292
 DB 2049 TAAAGATGATGAGATGCTCTTAAGACCAAGAAATGTCATATGACATGACACCTAG 2108
 QY 2223 CCAGAAAGGCTGTGTCAGAGACTTCATGAGCTCAACCAACCCCAATCTCAACAGAT 2352
 DB 2109 CCAGAAAGGCTGTGTCAGAGACTTCATGAGCTCAACCAACCCCAATCTCAACAGAT 2168
 QY 2353 GGATGAGTCTGTGTAACAAGTTTGGAAAAAGGACCGAAAAACAACGTGCGCAAAAGGCCAGG 2412
 DB 2169 GGATGAGTCTGTGTAACAAGTTTGGAAAAAGGACCGAAAAACAACGTGCGCAAAAGGCCAGG 2228
 QY 2413 ACATACCCCGAGGAAACAGTAACAATGGAAGCACTTACAGAAATTAAGAAAGTGAAGA 2472
 DB 2229 ACATACCCCGAGGAAACAGTAACAATGGAAGCACTTACAGAAATTAAGAAAGTGAAGA 2288
 QY 2473 CAGAGGACCCAGCAATTTGAGAGGGCAACCAAGAGTGTGAGCTGATTAACCTCTAGA 2532
 DB 2289 CAGAGGACCCAGCAATTTGAGAGGGCAACCAAGAGTGTGAGCTGATTAACCTCTAGA 2348
 QY 2533 AACCTCAAAACAAGTAAGAACTTGCCTAGACAATACTGAAAAACAATGCAATATACAT 2592
 DB 2349 AACCTCAAAACAAGTAAGAACTTGCCTAGACAATACTGAAAAACAATGCAATATACAT 2408
 QY 2553 GAACCTTTTTCATGAGCATTAATGTCATGTTTCAATGCTGGAAATTCAGCTGACTTCA 2652
 DB 2409 GAACCTTTTTCATGAGCATTAATGTCATGTTTCAATGCTGGAAATTCAGCTGACTTCA 2468
 QY 2653 CCAATTAATAATTAATCCATGAGTAACCTTCTTAATAGGCTTTT 2700
 DB 2469 CCAATTAATAATTAATCCATGAGTAACCTTCTTAATAGGCTTTT 2516

RESULT 4
 AA087442
 ID AA087442 standard; cDNA, 2601 BP.

AC AA087442;
 AC 25-MAR-2003 (updated)
 DT 21-NOV-1995 (first entry)
 XX Human semaphorin III cDNA.
 DE Semaphorin; grasshopper; human; vaccinia virus; Drosophila; Tribolium;
 KW variola major virus; smallpox; semaphorin receptor binding activity;
 KW modulation; nerve cell growth; immune response; viral pathogenesis;
 KW neurological disease; neuro-regeneration; oncological infection; dr.

OS Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 16..2331
 FT /tag= a
 FT /product= human semaphorin III

XX W09507706-A1.

XX 23-MAR-1995.

XX 13-SEP-1994; 94WO-US10151.

XX 13-SEP-1993; 93US-0121713.

XX (REGC) UNIV CALIFORNIA.

XX Bentley DR, Goodman CS, Kolodkin AL, Matthes D,

XX O'Connor T;

DR WP1; 1995-131177/17.
 DR P-PSDB; AAR71380.
 XX New class of semaphorin peptide(s) and polypeptide(s) - are
 PT potent modulators of nerve cell growth and regeneration
 PS Example 2; Page 60-63; 101pp; English.

CC The sequence of the cDNA encoding the human semaphorin III protein.
 CC The proteins encoded by the grasshopper semaphorin I (AA087441), human
 CC semaphorin III, vaccinia virus semaphorin IV (AA087443), Drosophila
 CC semaphorin I and II (AA087444-5), Tribolium semaphorin I (AA087446) or
 CC variola major (smallpox) virus semaphorin IV (AA087447) genes were used
 CC to generate a series of peptides (AAR70370-R70418), which retain
 CC semaphorin receptor binding activity. The semaphorin derived or
 CC semaphorin receptor derived peptides are potent modulators of nerve cell
 CC growth, immune responsiveness and viral pathogenesis. They can be used
 CC in diagnosis and treatment of neurological disease and
 CC neuro-regeneration, immune modulation and diagnosis and treatment of
 CC viral and oncological infection and diseases.
 CC (Updated on 25-MAR-2003 to correct PN field.)

SO Sequence 2601 BP; 809 A; 533 C; 593 G; 666 T; 0 other;

Query Match 92.6%; Score 2508; DB 16; Length 2601;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 CTGACAGATGGGCTGTGTTAACTAGAGATTGCTGCTTTCTGGGAGATTAATTACACGC 252
 DB 9 CTGACAGATGGGCTGTGTTAACTAGAGATTGCTGCTTTCTGGGAGATTAATTACACGC 68
 QY 253 AAGGCAACATCATGAGATGGGAGAAACAATGSCCAAGGCTGAATTAATCTCAAGA 312
 DB 69 AAGGCAACATCATGAGATGGGAGAAACAATGSCCAAGGCTGAATTAATCTCAAGA 128
 QY 313 AATGTGAAATCCAAACATGATGATCACTTCAATGCTGTGGCAACAGCTCCAGTTATCA 372
 DB 129 AATGTGAAATCCAAACATGATGATCACTTCAATGCTGTGGCAACAGCTCCAGTTATCA 188
 QY 373 TACCTTCCTTTGGATGAGAAACGAGTAGGCTGTATGTTGAGCAAAAGATCAATATT 432
 DB 189 TACCTTCCTTTGGATGAGAAACGAGTAGGCTGTATGTTGAGCAAAAGATCAATATT 248
 QY 433 TTCAATTCGACCTGTTAATCAAGATTTTCAAAAGATGTTGGCCAGATCTTACAC 492
 DB 249 TTCAATTCGACCTGTTAATCAAGATTTTCAAAAGATGTTGGCCAGATCTTACAC 308
 QY 493 CAGAAGAGATGAAGCAAGTGGCTGGAAAAAGACATCCTGAAAGATGCTAAATTTTCAAT 552
 DB 309 CAGAAGAGATGAAGCAAGTGGCTGGAAAAAGACATCCTGAAAGATGCTAAATTTTCAAT 368
 QY 553 CAAGTAATTAGGATTAATATCAAGTCACTTGTAGCTGTGGAAACGGGGCTTTTCA 612
 DB 369 CAAGTAATTAGGATTAATATCAAGTCACTTGTAGCTGTGGAAACGGGGCTTTTCA 428
 QY 613 TTCAATTTGACCTCAATGAAATTTGAGATCATCCGAGGACATATTTTAAAGCTGGA 672
 DB 429 TTCAATTTGACCTCAATGAAATTTGAGATCATCCGAGGACATATTTTAAAGCTGGA 488
 QY 673 GAACCTCAATTTGAAAAACGGCCGTGGGAAAGAGTCCATATGACCTTAAGCTGTACAGC 732
 DB 489 GAACCTCAATTTGAAAAACGGCCGTGGGAAAGAGTCCATATGACCTTAAGCTGTACAGC 548
 QY 733 ATCCCTTTAATAGATGAGATTAATCTGTGAATGACGTGAGCTGATTTTATGGGCGAGA 792
 DB 549 ATCCCTTTAATAGATGAGATTAATCTGTGAATGACGTGAGCTGATTTTATGGGCGAGA 608
 QY 793 CTTTGCATCTTCGGAATCTCTGGGACCAACCAATCAAGACAGACAGCATGATTC 852
 DB 609 CTTTGCATCTTCGGAATCTCTGGGACCAACCAATCAAGACAGCATGATTC 668
 QY 853 CAGGTGCTCAATGATCAAAAGTTCAATTAGTCCCACTCATCTCAGAGAGTGAATCC 912

Db		1749	CAGCCCTGAAAGAGAATCAATCTATNGGTGTAGAGAAATAGTAGCATTTTTGGAAATCGAG	1808		
Qy		1993	TCCGAAGTCGCAGAGAGCGCTGGTCTTAATGGCAAATTCOAAGSCGAAATGAGACGAAA	2052		
Db		1809	TCCGAAGTCGCAGAGAGCGCTGGTCTAATGGCAAATTCOAAGSCGAAATGAGACGAAA	1868		
Qy		2053	AGAAAGATCTCAGATGATGATCATATCATCAGGACAGATCAAGGCCCTTGCTGATAGTAG	2112		
Db		1869	AGAAAGATCTCAGATGATGATCATATCATCAGGACAGATCAAGGCCCTTGCTGATAGTAG	1928		
Qy		2113	TCCTCAACAGAGAAGATTCAGGCAATTAACCTTCGCGATCGCGTGGAMATGGGTTCATACA	2172		
Db		1929	TCCTCAACAGAGAAGATTCAGGCAATTAACCTTCGCGATCGCGTGGAMATGGGTTCATACA	1988		
Qy		2173	AACCTTCTTAAGGTAAACCCTGGAAATCATGACACAGAGCATTTGGAAAGACTTCTTCA	2232		
Db		1989	AACCTTCTTAAGGTAAACCCTGGAAATCATGACACAGAGCATTTGGAAAGACTTCTTCA	2048		
Qy		2233	TAAAGATGATCATGAGATGCGCTCTTAAGACCAAGAAATGTCCTCAATGACATGACACTAG	2292		
Db		2049	TAAAGATGATGATGAGATGCGCTCTTAAGACCAAGAAATGTCCTCAATGACATGACACTAG	2108		
Qy		2293	CCAGAAGGTCGTGTACAGAGACTTCATGACAGCTCATCAACACCCCAATCTCAACAGAT	2352		
Db		2109	CCAGAAGGTCGTGTACAGAGACTTCATGACAGCTCATCAACACCCCAATCTCAACAGAT	2168		
Qy		2353	GGATGAGTCTGTGAAACAAGTTGGAAAAAGGGAACGAAACACGTCGGCAAAAGCCAGG	2412		
Db		2169	GGATGAGTCTGTGAAACAAGTTGGAAAAAGGGAACGAAACACGTCGGCAAAAGCCAGG	2228		
Qy		2413	ACATATCCCACGAGGAACAGTAAACAATGGAAGCCTTCAAGAAAAATAAGAAAGGTAGAAA	2472		
Db		2229	ACATATCCCACGAGGAACAGTAAACAATGGAAGCCTTCAAGAAAAATAAGAAAGGTAGAAA	2288		
Qy		2473	CAGAGAGACCCACGAATTTTGAGAGGGGACCCAGGAGTGTCTGAGCTGCATTACCTCTAGA	2532		
Db		2289	CAGAGAGACCCACGAATTTTGAGAGGGGACCCAGGAGTGTCTGAGCTGCATTACCTCTAGA	2348		
Qy		2533	AACCTCAAAACAAGTAGAACTTGCTGTAGACAACTAACCTGAAAAACAATGCAATATACAT	2592		
Db		2349	AACCTCAAAACAAGTAGAACTTGCTGTAGACAACTAACCTGAAAAACAATGCAATATACAT	2408		
Qy		2593	GAACTTTTTTCATGGCATTAATGTGGAGTTTACAATGTGGGAAATCAGCTGAGTTCCA	2652		
Db		2409	GAACTTTTTTCATGGCATTAATGTGGAGTTTACAATGTGGGAAATCAGCTGAGTTCCA	2468		
Qy		2653	CCAAATTAATAATTAATTCATGAGTAACCTTCCAAATAGGCTTTTTTT	2700		
Db		2469	CCAAATTAATAATTAATTCATGAGTAACCTTCCAAATAGGCTTTTTTT	2516		
RESULT 5						
ID	AAO92331	standard; cDNA: 1481 BP.				
AC	AAO92331;					
XX						
DT	01-NOV-1995	(first entry)				
XX						
DE	Human collapsin cDNA.					
XX						
KW	Collapsin; antibody; therapy; ds.					
XX						
OS	Homo sapiens.					
XX						
FH	Key	Location/Qualifiers				
FT	CDS	50..1480				
XX		/tag= a				
PN	US5416197-A.					
XX						
PD	16-MAY-1995.					

PF 15-OCT-1993; 93US-0136922.
 XX 15-OCT-1993; 93US-0136922.
 XX (UNPE-) UNIV PENNSYLVANIA.
 PA Luo Y, Raper JA;
 XX MPI, 1995-193478/25.
 DR P-PSDB; AAR74175.
 XX
 PT New antibody to human collapsin - used to inhibit the activity of
 PT collapsin, to induce neurite out-growth and to treat individuals with
 PT nerve damage.
 PS Disclosure: Columns 11-16; 11pp; English.
 CC Human collapsin and its encoding nucleic acid may be used to
 CC identify agents which modulate the ability of human collapsin to
 CC collapse the growth cone of neurons. An antibody capable of
 CC specifically binding at least a portion of the collapsin protein can
 CC be used to purify human collapsin and to inhibit the activity of the
 CC protein. It can be used to induce neurite outgrowth by neuronal
 CC cells and to treat individuals suffering from nerve damage.
 CC
 SQ Sequence 1481 BP; 454 A; 299 C; 345 G; 383 T; 0 other;
 Query Match 46.3%; Score 1253; DB 16; Length 1481;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 103; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 904 TGACATCTCTGAAGATGACAAAGTACTTTCTTCCTGGAATGCAATAGATGAGA 963
 DB 58 TGACATCTCTGAAGATGACAAAGTACTTTCTTCCTGGAATGCAATAGATGAGA 117
 QY 964 ACACTCTGGAAGGCTACTACGCTGAGATGAGATGCAATGCAAGATGACTTTGAGG 1023
 DB 118 ACACTCTGGAAGGCTACTACGCTGAGATGAGATGCAATGCAAGATGACTTTGAGG 177
 QY 1024 GCACAGAGTCTGGTGAATTAATGAGCAAACTTCCTCAAGTCTGATTTGCTCAGT 1083
 DB 178 GCACAGAGTCTGGTGAATTAATGAGCAAACTTCCTCAAGTCTGATTTGCTCAGT 237
 QY 1084 GCCAGGTCCAATGCGATTGACACTCATTTGATGATGACAGATGATTCCTATGAA 1143
 DB 238 GCCAGGTCCAATGCGATTGACACTCATTTGATGATGACAGATGATTCCTATGAA 297
 QY 1144 CTTTAAAGATCTTAAAAATCCAGTTGATATGAGATGTTTACGACTTCAGTAACATTTT 1203
 DB 298 CTTTAAAGATCTTAAAAATCCAGTTGATATGAGATGTTTACGACTTCAGTAACATTTT 357
 QY 1204 CAAGGATCAGCCGTGTATATGATAGATGATGAGAGGCTTCCTTGGTCC 1263
 DB 358 CAAGGATCAGCCGTGTATATGATAGATGATGAGAGGCTTCCTTGGTCC 417
 QY 1264 ATATGCCCAAGGATGAGACCAACTATCAATGGGTGCTTATCAAGGAAGTCCCTTA 1323
 DB 418 ATATGCCCAAGGATGAGACCAACTATCAATGGGTGCTTATCAAGGAAGTCCCTTA 477
 QY 1324 TCCACGGCCAGGAATTTGCCAGCAAAATTTGGTGGTTTGACTTACAAAGAGACT 1383
 DB 478 TCCACGGCCAGGAATTTGCCAGCAAAATTTGGTGGTTTGACTTACAAAGAGACT 537
 QY 1384 TCCATGATATGTTAAACCTTTGCAAGAGATCATCAACCATGATCAATCCAGTTCCTC 1443
 DB 538 TCCATGATATGTTAAACCTTTGCAAGAGATCATCAACCATGATCAATCCAGTTCCTC 597
 QY 1444 TATGAACATCGCCCAATAGTATCAAAACGATGTAATTAATCAATTAACAATTTGT 1503
 DB 598 TATGAACATCGCCCAATAGTATCAAAACGATGTAATTAATTAACAATTTGT 657
 QY 1504 CGTAGACCGAGTGTATGAGAAAGATGACAGTATGATGTTATCGAAGCAGATGT 1563

DB 658 CGTAGACCGAGTGTATGAGAAATGACAGTATGATGTTATCGAAGCAGATGT 717
 QY 1564 TGGGACCGTCTTAAAGATGTTCAATTCCTAAGAGACTGTGATGATTAGAGAGT 1623
 DB 718 TGGGACCGTCTTAAAGATGTTCAATTCCTAAGAGACTGTGATGATTAGAGAGT 777
 QY 1624 TCTGCTGGAAGAAATGACAGTTTTCGGGAACCGACTGCTATTTTCAAGCAATGAGCTTTC 1683
 DB 778 TCTGCTGGAAGAAATGACAGTTTTCGGGAACCGACTGCTATTTTCAAGCAATGAGCTTTC 837
 QY 1684 CACTAAGCAGCAACACTATATATTTGTTCAAGCGCTGGGGTTCCTCCAGCTCCTTTACA 1743
 DB 838 CACTAAGCAGCAACACTATATATTTGTTCAAGCGCTGGGGTTCCTCCAGCTCCTTTACA 897
 QY 1744 CCGGTGTATATTAACGGGAAGCGTGTGCTGAGTGTGCTGCTGCGCCGAGACCTTTACTG 1803
 DB 898 CCGGTGTATATTAACGGGAAGCGTGTGCTGAGTGTGCTGCTGCGCCGAGACCTTTACTG 957
 QY 1804 TGCTTGGGATGTTCTGCAATGTTCTGCTATTTTCCACTGCAAAAGAGACGCAAGACG 1863
 DB 958 TGCTTGGGATGTTCTGCAATGTTCTGCTATTTTCCACTGCAAAAGAGACGCAAGACG 1917
 QY 1864 ACAAGATATTAAGAAATGAGACCCACTGACTCACTGTTCACTTACCACTATGATATCA 1923
 DB 1918 ACAAGATATTAAGAAATGAGACCCACTGACTCACTGTTCACTTACCACTATGATATCA 1977
 QY 1924 CCATGGCCACAGCCCTGAAGAGATCATCTATGATGATGAGAAATGATGACATTTT 1983
 DB 1078 CCATGGCCACAGCCCTGAAGAGATCATCTATGATGATGAGAAATGATGACATTTT 1137
 QY 1984 GGAATGAGTCCGAGTCCGAGAGGCGCTGCTATTTGCAATTCAGAGCCGAAATGA 2043
 DB 1138 GGAATGAGTCCGAGTCCGAGAGGCGCTGCTATTTGCAATTCAGAGCCGAAATGA 1197
 QY 2044 AGACGGAAGAGAGATCAGATGATGATCATATCATCAGACAGATCAAGCCCTTCT 2103
 DB 1198 AGACGGAAGAGAGATCAGATGATGATCATATCATCAGACAGATCAAGCCCTTCT 1257
 QY 2104 GCTAGTATCTCAACAGAAAGATTCAGGCAATTCCTGCGCATCGGTGGAACATGG 2163
 DB 1258 GCTAGTATCTCAACAGAAAGATTCAGGCAATTCCTGCGCATCGGTGGAACATGG 1317
 QY 2164 GTTCATACAACTCTTCTTAAGTACCTGGAAGTATTGACA 2207
 DB 1318 GTTCATACAACTCTTCTTAAGTACCTGGAAGTATTGACA 1361

RESULT 6
 ABS45213
 ID ABS45213 standard; DNA; 456 BP.
 XX
 AC ABS45213;
 XX
 DT 25-FEB-2003 (first entry)
 XX
 DE Human liver single exon probe, SEQ ID No 20203.
 XX
 KW Human; single exon nucleic acid probe; liver; cirrhosis;
 KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
 KW coronary heart disease; ss.
 OS Homo sapiens.
 XX
 EN WO200157273-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00664.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2001-488898/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human adult liver -
 PS Claim 4; SEQ ID No 20203; 658bp; English.
 XX
 CC The invention relates to a single exon nucleic acid probe (SENP) (1) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridizes at high
 CC stringency to a nucleic acid molecule expressed in the human adult
 CC liver. (1) may be used for predicting, measuring and displaying gene
 CC expression in samples derived from human adult liver. The genes
 CC identified may be involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
 CC is associated with coronary heart disease. ABS25011-ABS51005 represent
 CC human liver single exon nucleic acid probes of the invention.
 CC Note: The sequence information for this patent does not appear in the
 CC printed specification but was obtained in electronic format directly
 CC from WFO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 456 BP; 161 A; 98 C; 108 G; 89 T; 0 other;
 Query Match 15.0%; Score 405; DB 23; Length 456;
 Best Local Similarity 99.8%; Pred. No. 2.6e-189;
 Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2060 ATCAGAGTGCATGATCATATCATCAGACAGATCAAGCCCTTCTGCTACGTAGCTTACAA 2119
 DB 1 ATCAGAGTGCATGATCATATCATCAGACAGATCAAGCCCTTCTGCTACGTAGCTTACAA 60
 QY 2120 CAGAGGATTCAGGCAATACCTCTGCGCATGCGGTGGAACATGGGTCTCATACAACTCTT 2179
 DB 61 CAGAGGATTCAGGCAATACCTCTGCGCATGCGGTGGAACATGGGTCTCATACAACTCTT 120
 QY 2180 CTTAAGGTAACTCTGGAAGTCAATTCACACAGAGCATTTTGGAGAACTTCTTCAATAAGAT 2239
 DB 121 CTTAAGGTAACTCTGGAAGTCAATTCACACAGAGCATTTTGGAGAACTTCTTCAATAAGAT 180
 QY 2240 GATGATGAGATGGCTCTTAAGACCAAAAGATGTCCAATAGCATGACACTTACCCAGAAG 2299
 DB 181 GATGATGAGATGGCTCTTAAGACCAAAAGATGTCCAATAGCATGACACTTACCCAGAAG 240
 QY 2300 GTCCTGTACAGAGCTTCTATGACGCTCATCAACCCCAATTCATCAACGATGATGAG 2359
 DB 241 GTCCTGTACAGAGCTTCTATGACGCTCATCAACCCCAATTCATCAACGATGATGAG 300
 QY 2360 TTCTGTGAACAAGTTTGAAGAAAGGACCAAAACCAAGTGGCAAGGCGAGACATACC 2419
 DB 301 TTCTGTGAACAAGTTTGAAGAAAGGACCAAAACCAAGTGGCAAGGCGAGACATACC 360
 QY 2420 CCAGGAGACAGTACCAATGAAAGCACTTACCAAGAAATTAAGAAAGTGAAGAAACAGGAGG 2479
 DB 361 CCAGGAGACAGTACCAATGAAAGCACTTACCAAGAAATTAAGAAAGTGAAGAAACAGGAGG 420
 QY 2480 ACCCAAGATTTGAGAGGGCCACCGAGAGTGTCTGA 2515
 DB 421 ACCCAAGATTTGAGAGGGCCACCGAGAGTGTCTGA 456
 RESULT 7
 ABS19795
 ID ABS19795 standard; DNA; 456 BP.
 XX

AC ABS19795;
 XX 19-AUG-2002 (first entry)
 DT
 XX
 DE Human genome-derived single exon probe ORF from lung SEQ ID No 19786.
 XX
 KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease; open reading frame; ORF.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 PP 30-JAN-2001; 2001WO-US00665.
 XX
 PR 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 DR WPI; 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 PS Claim 4; SEQ ID No 19786; 634bp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridize at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease.

CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemostatic disorder, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karsagen syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe open reading frame of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIP at
 CC ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 456 BP; 161 A; 98 C; 108 G; 89 T; 0 other;

Query Match 15.0%; Score 405; DB 24; Length 456;
 Best Local Similarity 99.8%; Pred. No. 2.6e-189;
 Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2060 ATCAGAGTGGATGATCATATCTATCAGACAGATCAAGGCTTCTGCTACGATGCTACAA 2119
 DB 1 ATCAGAGTGGATGATCATATCTATCAGACAGATCAAGGCTTCTGCTACGATGCTACAA 60
 QY 2120 CAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2179
 DB 61 CAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 QY 2180 CTTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2239
 DB 121 CTTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 QY 2240 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2299
 DB 181 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 240
 QY 2300 GTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2359
 DB 241 GTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 QY 2360 TTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2419
 DB 301 TTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
 QY 2420 CAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2479
 DB 361 CAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 420
 QY 2480 ACCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2515
 DB 421 ACCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 456

RESULT 8
 AA129182
 ID AA129182 standard; cDNA, 354 BP.

AC AA129182;

DT 12-OCT-2001 (first entry)

DE Colon tumour related determined cDNA sequence for clone R0093.F06.

KW Human; immunotherapy; diagnosis; colon cancer; colon tumour;
 KW immunogenic; gene therapy; vaccine; colonic cancer; ss.

OS Homo sapiens.

XX WO200149716-A2.

PD 12-JUL-2001.

PF 29-DEC-2000; 2000WO-US35596.

XX 30-DEC-1999; 99US-0476296.

PR 10-JAN-2000; 2000US-0480321.

PR 15-FEB-2000; 2000US-0504629.
 PR 06-MAR-2000; 2000US-0519444.
 PR 19-MAY-2000; 2000US-0575251.
 PR 29-JUN-2000; 2000US-0609448.
 PR 28-AUG-2000; 2000US-0649811.

PA (CORI-) CORIXA CORP.

PI Xu J, Lodes MJ, Secret H, Benson DR, Meagher MJ, Stolk JA;

PI King GE, Wang T, Jiang Y;

DR WPI; 2001-441847/47.

CC Colon tumor associated proteins and nucleic acids useful for the
 CC prevention, diagnosis and treatment of colonic cancer -

PS Claim 2; Page 342; 472pp; English.

The present invention describes colon tumour associated proteins (I) and
 the polynucleotides (II) that encode them. (I) have cytostatic activity.
 (I) and (II) can be used in gene therapy and vaccine production. (I) and
 (II) may be used in the prevention, diagnosis and treatment of diseases
 associated with inappropriate colon tumour associated protein (TCAP)
 expression, such as colonic cancer. For example, (I) and (II) may be
 used to treat disorders associated with decreased expression by
 rectifying mutations or deletions in a patient's genome that affect the
 activity of TCAPs by expressing inactive proteins or to supplement the
 patients own production of them. Additionally, (II) may be used to
 produce the TCAP proteins, by inserting the nucleic acid into a host
 cell culturing the cell to express the protein. (II) and its
 complementary sequences may also be used as DNA probes in diagnostic
 PCR and hybridisation assays to detect and
 quantify the presence of similar nucleic acids in samples, and
 therefore which patients may be in need of restorative therapy. (I) may
 also be used as antigens in the production of antibodies against TCAPs
 and in assays to identify modulators of TCAP expression and activity.
 Anti-(I) antibodies and antagonists may also be used to down regulate
 TCAP expression and activity. The anti-(I) antibodies may also be used
 as diagnostic agents for detecting the presence of TCAPs in samples
 (e.g. by enzyme linked immunosorbent assay (ELISA)). AA128460 to AA129512
 CC and AA124494 to AA124533 represent nucleotide and amino acid sequences
 CC given in the exemplification of the present invention.

Sequence 354 BP; 120 A; 82 C; 79 G; 72 T; 1 other;

Query Match 11.2%; Score 303; DB 22; Length 354;
 Best Local Similarity 99.7%; Pred. No. 6.2e-139;
 Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2098 CTTCTGCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2157
 DB 1 CTTCTGCTAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 60
 QY 2158 ACATGGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2217
 DB 61 ACATGGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 120
 QY 2218 GGAAGAACTTCTTATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2277
 DB 121 GGAAGAACTTCTTATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 QY 2278 TAGCATGACCTGACGAGAGAGGCTGTGACAGAGCTTCAATGAGCTCATCAACCAACC 2337
 DB 181 TAGCATGACCTGACGAGAGAGGCTGTGACAGAGCTTCAATGAGCTCATCAACCAACC 240
 QY 2338 CAATCTCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2397
 DB 241 CAATCTCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
 QY 2398 TCGGCAAGGCGAGAGATATCCCGAGGAACAGTAACAAATGGAAGCACTTACA 2451
 DB 301 TCGGCAAGGCGAGAGATATCCCGAGGAACAGTAACAAATGGAAGCACTTACA 354

RESULT 9
ABZ33368
ID ABZ33368 standard; cDNA; 354 BP.
XX
AC ABZ33368;
XX
DT 30-JAN-2003 (first entry)
XX
DE Human colon tumour cDNA for clone R0093; F06 SEQ ID NO:736.
XX
KM Human; colon cancer; colon tumour; immunotherapy; diagnosis; cancer;
KW tumour; immune response; immunostimulant; cytostatic; vaccine;
KW gene; ss.
XX
OS Homo sapiens.
XX
PN WO200283070-A2.
XX
PD 24-OCT-2002.
XX
PF 09-APR-2002; 2002WO-US11475.
XX
PR 10-APR-2001; 2001US-0833263.
XX
PR 03-AUG-2001; 2001US-0922217.
XX
PR 19-DEC-2001; 2001US-0025380.
XX
PA (CORI-) CORIXA CORP.
XX
PI Xu J, Lodes MJ, Secrist H, Benson DR, Meagher MJ, Stolk JA,
PI Wang T, Jiang Y, Smith CL, King GE, Wang A, Clapper JD,
PI Skeiky YAW, Fanger GR, Vedvick TS, Carter D;
XX
DR WPI: 2003-067548/06.
XX
PT New polynucleotide, useful for the preparation of a composition for
PT stimulating an immune response against, or treating, cancer
XX
PS Disclosure; Page 343; 537pp; English.
XX
CC The present invention describes compounds (1) for the immunotherapy and
CC diagnosis of colon cancer. Also described: (1) a method for detecting
CC the presence of cancer in a patient; (2) a method for stimulating and/or
CC expanding T cells specific for a tumour protein; (3) an isolated T cell
CC population comprising T cells prepared by the method of (2); (4) a method
CC for stimulating an immune response in a patient; (5) a method for
CC treating cancer in a patient; and (6) a method for inhibiting the
CC development of cancer in a patient. (1) have immunostimulant and
CC cytostatic activities and can be used in vaccines. ABZ32646 to ABZ33725
CC and ABP55343 to ABP55391 represent human colon cancer/tumour related
CC sequences used in the exemplification of the present invention.
XX
SQ Sequence 354 BP; 120 A; 82 C; 79 G; 72 T; 1 other;
XX
Query Match 11.2%; Score 303; DB 25; Length 354;
Best Local Similarity 99.7%; Pred. No. 6,2e-139;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
QY 2098 CCTTGTGCTAGCTACTTCAACAGAGGATTGAGGCAATTCCTGCGATGCGGTGA 2157
DB 1 CTTTGTGCTAGCTAGCTTCAACAGAGGATTGAGGCAATTCCTGCGATGCGGTGA 60
QY 2158 ACATGGTTTCATACAACTCTTCTTAAGGTAAACCTGGAAGTCAATGACAGAGATT 2217
DB 61 ACATGGTTTCATACAACTCTTCTTAAGGTAAACCTGGAAGTCAATGACAGAGATT 120
QY 2218 GGAAGAACTTCTTCAATAAAGATGATGAGATGGCTTAAAGCAAGAAATGTCGA 2277
DB 121 GGAAGAACTTCTTCAATAAAGATGATGAGATGGCTTAAAGCAAGAAATGTCGA 180
QY 2278 TAGCATGACACTAGCGAAGGCTGCTGAGACAGACTTCATGACAGCTATTAACACCC 2337
DB 181 TAGCATGACACTAGCGAAGGCTGCTGAGACAGACTTCATGACAGCTATTAACACCC 240

QY 2338 CAATCTCAACAGATGATGAGTTCTGTGAACAAGTTTGAAGAAAGGACCCGAAACAAAG 2397
DB 241 CAATCTCAACAGATGATGAGTTCTGTGAACAAGTTTGAAGAAAGGACCCGAAACAAAG 300
QY 2398 TCGGCAAAAGCCAGGACATACCCAGGAGACATGATACAAATGAGACACTTACA 2451
DB 301 TCGGCAAAAGCCAGGACATACCCAGGAGACATGATACAAATGAGACACTTACA 354
XX
RESULT 10
ABZ32213
ID ABZ32213 standard; DNA; 496 BP.
XX
AC ABZ32213;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID NO 7203.
XX
KM Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinemia; hyperlipidaemia; hypercholesterolemia;
KW coronary heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00664.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI: 2001-488898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analysing gene expression in human adult liver
XX
PS Claim 1; SEQ ID NO 7203; 658bp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (1) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult
CC liver. (1) may be used for predicting, measuring and displaying gene
CC expression in samples derived from human adult liver. The genes
CC identified may be involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinemia, hyperlipidaemia and hypercholesterolemia which
CC is associated with coronary heart disease. ABZ5011-ABZ51005 represent
CC human liver single exon nucleic acid probes of the invention.
CC Note: The sequence information for this patent does not appear in the
CC printed specification but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pcr_sequences.
XX
SQ Sequence 496 BP; 164 A; 101 C; 92 G; 139 T; 0 other;
XX
Query Match 10.8%; Score 292; DB 23; Length 496;
Best Local Similarity 100.0%; Pred. No. 1,7e-133;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 2058 AGATCAGATGATGATCATATCATCAAGACGATCAAGGCTTCTGCTAGTACTTAC 2117

```

160 AGATCAGGTGATGATCATATCATCATGAGACGATCAAGGCGCTTGTGCTACGCTTAC
219
2118 AACAGAGGATTCAGGCAATTACCTCTGCGCATGCGGTGGAGACATGGGTTTCATACAACTC
2177
220 AACAGAGGATTCAGGCAATTACCTCTGCGCATGCGGTGGAGACATGGGTTTCATACAACTC
279
2178 TTCTTAAGGTAACTCTGGAAGTATGACACAGAGCATTTGGAGAACTTCTTCAATAAG
2237
280 TTCTTAAGGTAACTCTGGAAGTATGACACAGAGCATTTGGAGAACTTCTTCAATAAG
339
2238 ATGATGATGAGATGGCTCTTAAGACCAAGAAATGTCCATATGATGACCTTACCCAGA
2297
340 ATGATGATGAGATGGCTCTTAAGACCAAGAAATGTCCATATGATGACCTTACCCAGA
399
2238 AGGTCTGTACAGAGACTTTCATGACGCTCATGACACCCCAATCTCAACAC
2349
400 AGGTCTGTACAGAGACTTTCATGACGCTCATGACACCCCAATCTCAACAC
451

```

RESULT 11

ABS07290

ID ABS07290 standard; DNA; 496 BP.

AC ABS07290;

DT 19-AUG-2002 (first entry)

DE Human genome-derived single exon probe from lung SEQ ID No 7281.

Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 Chronic obstructive pulmonary disease; interstitial lung disease;
 familial idiopathic pulmonary fibrosis; neurofibromatosis;
 tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
 pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 primary ciliary dyskinesia; pulmonary hypertension;
 hyaline membrane disease.

OS Homo sapiens.

PN WO200186003-A2.

PD 15-NOV-2001.

PE 30-JAN-2001; 2001WO-US00665.

PR 04-FEB-2000; 2000US-180312P.

PR 26-MAY-2000; 2000US-207456P.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-234687P.

PR 27-SEP-2000; 2000US-236539P.

PR 04-OCT-2000; 2000GB-0024263.

PA (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

PS WPI; 2002-114183/15.

PT Spatially-addressable set of single exon nucleic acid probes, used to

PT measure gene expression in human lung samples -

PS Claim 1; SEQ ID No 7281; 634bp; English.

CC The invention relates to a spatially-addressable set of single exon

CC nucleic acid probes for measuring gene expression in a sample derived

CC from human lung comprising single exon nucleic acid probes having one of

CC 12614 nucleic acid sequences mentioned in the specification, or their

CC complements or the 12387 open reading frames derived from the 12614

CC probes. Also included are a microarray comprising the novel set of

CC probes; the novel set of probes which hybridize at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridization of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridization to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 496 BP; 164 A; 101 C; 92 G; 139 T; 0 other;

Query Match 10.8%; Score 292; DB 24; Length 496;

Best Local Similarity 100.0%; Pred. No. 1.7e-113; Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 2058 AGATCAGGTGATGATCATATCATCATGAGACGATCAAGGCGCTTGTGCTACGCTTAC 2117
DB 160 AGATCAGGTGATGATCATATCATCATGAGACGATCAAGGCGCTTGTGCTACGCTTAC 219
QY 2118 AACAGAGGATTCAGGCAATTACCTCTGCGCATGCGGTGGAGACATGGGTTTCATACAACTC 2177
DB 220 AACAGAGGATTCAGGCAATTACCTCTGCGCATGCGGTGGAGACATGGGTTTCATACAACTC 279
QY 2178 TTCTTAAGGTAACTCTGGAAGTATGACACAGAGCATTTGGAGAACTTCTTCAATAAG 2237
DB 280 TTCTTAAGGTAACTCTGGAAGTATGACACAGAGCATTTGGAGAACTTCTTCAATAAG 339
QY 2238 ATGATGATGAGATGGCTCTTAAGACCAAGAAATGTCCATATGATGACCTTACCCAGA 2297
DB 340 ATGATGATGAGATGGCTCTTAAGACCAAGAAATGTCCATATGATGACCTTACCCAGA 399
QY 2298 AGGTCTGTACAGAGACTTTCATGACGCTCATGACACCCCAATCTCAACAC 2349
DB 400 AGGTCTGTACAGAGACTTTCATGACGCTCATGACACCCCAATCTCAACAC 451

```

RESULT 12

ABA71205/c

ID ABA71205 standard; DNA; 172 BP.

AC ABA71205;

DT 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe. #19510.

KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX OS Homo sapiens.
 XX PD WO200157277-A2.
 XX PN 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US00669.
 XX PR 04-FEB-2000; 2000US-0180312.
 XX PR 26-MAY-2000; 2000US-0207456.
 XX PR 30-JUN-2000; 2000US-0608408.
 XX PR 03-AUG-2000; 2000US-0632366.
 XX PR 21-SEP-2000; 2000US-0234687.
 XX PR 27-SEP-2000; 2000US-0236359.
 XX PR 04-OCT-2000; 2000GB-0024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX DR WPI; 2001-483447/52.
 XX PT Human genome-derived single exon nucleic acid probes useful for
 XX PT analyzing gene expression in human fetal liver -
 XX PS Claim 4; SEQ ID NO 19510; 639pp + sequence listing; English.
 XX CC The invention relates to a single exon nucleic acid probe for
 XX CC measuring human gene expression in a sample derived from human fetal
 XX CC liver. The single exon nucleic acid probes may be used for predicting,
 XX CC measuring and displaying gene expression in samples derived from human
 XX CC fetal liver. The present sequence is a single exon nucleic acid
 XX CC probe of the invention.
 XX CC Note: The sequence data for this patent did not form part of the
 XX CC printed specification, but was obtained in electronic format directly
 XX CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 172 BP; 54 A; 37 C; 32 G; 49 T; 0 other;
 XX

Query Match 5.9%; Score 160; DB 22; Length 172;
 Best Local Similarity 100.0%; Pred. No. 2.6e-68;
 Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCACATGATGATCACTTTCATGCTTGCCCAAGCTCCAGTTA 369
 DB 160 AGAATGTTGGAATCCACATGATGATCACTTTCATGCTTGCCCAAGCTCCAGTTA 101
 QY 370 TCATACCTTCTTTGGATGAGAACGAGTAGGCTGTATGTTGAGCAAGATCAGAT 429
 DB 100 TCATACCTTCTTTGGATGAGAACGAGTAGGCTGTATGTTGAGCAAGATCAGAT 41
 QY 430 ATTTTCATTCGACCTGTTAATATCAAGATTTTCAAAAG 469
 DB 40 ATTTTCATTCGACCTGTTAATATCAAGATTTTCAAAAG 1

RESULT 13
 AAK19504/C
 ID AAK19504 standard; DNA; 172 BP.
 XX AAK19504;
 AC
 XX
 DT 05-NOV-2001 (first entry)
 XX
 DE Human brain expressed single exon probe SEQ ID NO: 19495.
 XX
 KW Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer; ss.
 XX
 OS Homo sapiens.
 XX

PN WO200157275-A2.
 XX
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US00667.
 XX PR 04-FEB-2000; 2000US-0180312.
 XX PR 26-MAY-2000; 2000US-0207456.
 XX PR 30-JUN-2000; 2000US-0608408.
 XX PR 03-AUG-2000; 2000US-0632366.
 XX PR 21-SEP-2000; 2000US-0234687.
 XX PR 27-SEP-2000; 2000US-0236359.
 XX PR 04-OCT-2000; 2000GB-0024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX DR WPI; 2001-483446/52.
 XX PT Single exon nucleic acid probes for analyzing gene expression in human
 XX PT brains -
 XX PS Example 4; SEQ ID NO: 19495; 650pp + Sequence Listing; English.
 XX CC The present invention provides a number of single exon nucleic acid
 XX CC probes which are derived from genomic sequences expressed in the human
 XX CC brain. They can be used to measure gene expression in brain cell samples,
 XX CC which may enable the diagnosis and improved treatment of nervous system
 XX CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 XX CC epilepsy and cancers. The present sequence is one of the probes of the
 XX CC invention.
 XX SQ Sequence 172 BP; 54 A; 37 C; 32 G; 49 T; 0 other;
 XX

Query Match 5.9%; Score 160; DB 22; Length 172;
 Best Local Similarity 100.0%; Pred. No. 2.6e-68;
 Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCACATGATGATCACTTTCATGCTTGCCCAAGCTCCAGTTA 369
 DB 160 AGAATGTTGGAATCCACATGATGATCACTTTCATGCTTGCCCAAGCTCCAGTTA 101
 QY 370 TCATACCTTCTTTGGATGAGAACGAGTAGGCTGTATGTTGAGCAAGATCAGAT 429
 DB 100 TCATACCTTCTTTGGATGAGAACGAGTAGGCTGTATGTTGAGCAAGATCAGAT 41
 QY 430 ATTTTCATTCGACCTGTTAATATCAAGATTTTCAAAAG 469
 DB 40 ATTTTCATTCGACCTGTTAATATCAAGATTTTCAAAAG 1

RESULT 14
 AAK45495/C
 ID AAK45495 standard; DNA; 172 BP.
 XX AAK45495;
 AC
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human bone marrow expressed single exon probe SEQ ID NO: 20052.
 XX
 KW Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukemia; lymphoma; myeloma; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157276-A2.
 XX
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US00668.
 XX

04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-488900/53.
 DR
 XX WPI; 2001-488900/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 XX
 PS Example 4; SEQ ID NO: 20052; 658bp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukemia and myeloma. The present sequence is one of
 CC the probes of the invention.
 CC
 SQ Sequence 172 BP; 54 A; 37 C; 32 G; 49 T; 0 other;

Query Match 5.9%; Score 160; DB 22; Length 172;
 Best Local Similarity 100.0%; Pred. No. 2.6e-68;
 Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGATCCACAATGATGATCACTTCAATGGCTTGGCCAAAGCTCCAGTTA 369
 DB 160 AGAATGTTGGATCCACAATGATGATCACTTCAATGGCTTGGCCAAAGCTCCAGTTA 101
 QY 370 TCATACCTCTCTTTGGATGAGAACGAGTAGGCTGTATGTTGGAGCAAGATCACAT 429
 DB 100 TCATACCTCTCTTTGGATGAGAACGAGTAGGCTGTATGTTGGAGCAAGATCACAT 41
 QY 430 ATTTTCATTGACCGGTGTTAATATCAAGATTTTCAAAAG 469
 DB 40 ATTTTCATTGACCGGTGTTAATATCAAGATTTTCAAAAG 1

RESULT 15
 AA151439/C
 ID AA151439 standard; DNA; 172 BP.
 XX
 AC AA151439;
 XX
 DT 17-OCT-2001 (first entry)
 XX
 DE Probe #20125 used to measure gene expression in human placenta sample.
 XX
 KW Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157272-A2.
 XX
 PD 09-AUG-2001.
 XX
 PE 30-JAN-2001; 2001WO-US00663.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.

XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-488900/53.
 DR
 XX WPI; 2001-488900/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 XX
 PS Claim 25; SEQ ID NO 20125; 654bp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENPs).
 CC The present sequence is one such probe. The probes are useful for
 CC producing a microarray for predicting, measuring and displaying gene
 CC expression in samples derived from human placenta. The probes are useful
 CC for antenatal diagnosis of human genetic disorders.
 CC
 SQ Sequence 172 BP; 54 A; 37 C; 32 G; 49 T; 0 other;

Query Match 5.9%; Score 160; DB 22; Length 172;
 Best Local Similarity 100.0%; Pred. No. 2.6e-68;
 Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGATCCACAATGATGATCACTTCAATGGCTTGGCCAAAGCTCCAGTTA 369
 DB 160 AGAATGTTGGATCCACAATGATGATCACTTCAATGGCTTGGCCAAAGCTCCAGTTA 101
 QY 370 TCATACCTCTCTTTGGATGAGAACGAGTAGGCTGTATGTTGGAGCAAGATCACAT 429
 DB 100 TCATACCTCTCTTTGGATGAGAACGAGTAGGCTGTATGTTGGAGCAAGATCACAT 41
 QY 430 ATTTTCATTGACCGGTGTTAATATCAAGATTTTCAAAAG 469
 DB 40 ATTTTCATTGACCGGTGTTAATATCAAGATTTTCAAAAG 1

RESULT 16
 ABS45181/C
 ID ABS45181 standard; DNA; 172 BP.
 XX
 AC ABS45181;
 XX
 DT 25-FEB-2003 (first entry)
 XX
 DE Human liver single exon probe, SEQ ID NO 20171.
 XX
 KW Human; single exon nucleic acid probe; liver; cirrhosis;
 KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
 KW coronary heart disease; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200157273-A2.
 XX
 PD 09-AUG-2001.
 XX
 PE 30-JAN-2001; 2001WO-US00664.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-488900/53.
 DR

PT Human genome-derived single exon nucleic acid probes useful for
 XX analysing gene expression in human adult liver -
 PS Claim 4; SEQ ID No 20171; 658bp; English.

CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult
 CC liver. (I) may be used for predicting, measuring and displaying gene
 CC expression in samples derived from human adult liver. The genes
 CC identified may be involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
 CC is associated with coronary heart disease. AB525011-AB551005 represent
 CC human liver single exon nucleic acid probes of the invention.
 CC Note: The sequence information for this patent does not appear in the
 CC printed specification but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 172 BP; 54 A; 37 C; 32 G; 49 T; 0 other;

QY Query Match 5.9%; Score 160; DB 23; Length 172;
 DB Best Local Similarity 100.0%; Pred. No. 2.6e-68;
 Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCAAATGATGATCACTTTCATGAGCTTGCCCAACAGCTCCAGTTA 369
 DB 160 AGAATGTTGGAATCCAAATGATGATCACTTTCATGAGCTTGCCCAACAGCTCCAGTTA 101

QY 370 TCATACCTTCTTTTGGATGAGAACGAGTGTATGTTGGAGCAAGGATCAAT 429
 DB 100 TCATACCTTCTTTTGGATGAGAACGAGTGTATGTTGGAGCAAGGATCAAT 41

QY 430 ATTTTCATTCGACCTGTATATATCAAGATTTTCAAAAG 469
 DB 40 ATTTTCATTCGACCTGTATATATCAAGATTTTCAAAAG 1

RESULT 17
 ABS19763/c
 ID ABS19763 standard; DNA; 172 BP.

AC ABS19763;
 XX
 XX
 DT 19-AUG-2002 (first entry)
 XX
 XX Human genome-derived single exon ORF from lung SEQ ID No 19754.
 DE Human genome-derived single exon ORF from lung SEQ ID No 19754.
 XX
 XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KM chronic obstructive pulmonary disease; interstitial lung disease;
 KM familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KM tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KM Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosterosiis;
 KM pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
 KM pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KM primary ciliary dyskinesia; pulmonary hypertension;
 KM hyaline membrane disease; open reading frame; ORF.

OS Homo sapiens.
 XX
 XX WO200186003-A2.
 XX
 PD 15-NOV-2001.
 XX
 XX 30-JAN-2001; 2001WO-US00665.
 PF
 XX 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.

PR 04-OCT-2000; 2000GB-0024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2002-114183/15.
 DR
 XX
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 CC Claim 4; SEQ ID No 19754; 634bp; English.

CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosterosiis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe open reading frame of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX
 SQ Sequence 172 BP; 54 A; 37 C; 32 G; 49 T; 0 other;

QY Query Match 5.9%; Score 160; DB 24; Length 172;
 DB Best Local Similarity 100.0%; Pred. No. 2.6e-68;
 Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCAAATGATGATCACTTTCATGAGCTTGCCCAACAGCTCCAGTTA 369
 DB 160 AGAATGTTGGAATCCAAATGATGATCACTTTCATGAGCTTGCCCAACAGCTCCAGTTA 101

QY 370 TCATACCTTCTTTTGGATGAGAACGAGTGTATGTTGGAGCAAGGATCAAT 429
 DB 100 TCATACCTTCTTTTGGATGAGAACGAGTGTATGTTGGAGCAAGGATCAAT 41

QY 430 ATTTTCATTCGACCTGTATATATCAAGATTTTCAAAAG 469
 DB 40 ATTTTCATTCGACCTGTATATATCAAGATTTTCAAAAG 1


```
RESULT 18
ID ABA58707/C
XX ABA58707 standard; DNA; 484 BP.
AC ABA58707;
XX
XX 01-FEB-2002 (first entry)
XX
XX Human foetal liver single exon nucleic acid probe #7012.
DE
XX Human foetal liver; gene expression; single exon nucleic acid probe; ss.
KM
XX Homo sapiens.
XX
XX WO200157277-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00669.
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX 26-MAY-2000; 2000US-0207456.
XX
XX 30-JUN-2000; 2000US-0608408.
XX
XX 03-AUG-2000; 2000US-0632366.
XX
XX 21-SEP-2000; 2000US-0234687.
XX
XX 27-SEP-2000; 2000US-0236359.
XX
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human fetal liver -
XX
XX Claim 1; SEQ ID NO 7012; 639pp + sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probes may be used for predicting,
XX measuring and displaying gene expression in samples derived from human
XX fetal liver. The present sequence is a single exon nucleic acid
XX probe of the invention.
XX
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 484 BP; 161 A; 77 C; 79 G; 167 T; 0 other;
SQ
Query Match 5.9%; Score 160; DB 22; Length 484;
Best Local Similarity 100.0%; Pred. No. 2.6e-68;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 310 AGAATGTTGGAATCCAAATGATGATCACTTCAATGCTTGGCCAAAGCTCCAGTTA 369
DB 458 AGAATGTTGGAATCCAAATGATGATCACTTCAATGCTTGGCCAAAGCTCCAGTTA 399
QY 370 TCATACCTTCTTTTGGATGAGAACGAGTAGAGCTGTATGTTGAGCAAAAGATCACAT 429
DB 398 TCATACCTTCTTTTGGATGAGAACGAGTAGAGCTGTATGTTGAGCAAAAGATCACAT 339
QY 430 ATTTTCATTGACCTGGTTAATATCAAGATTTTCAAAAG 469
DB 338 ATTTTCATTGACCTGGTTAATATCAAGATTTTCAAAAG 299
RESULT 19
ID AAK06842/C
XX AAK06842 standard; DNA; 484 BP.
XX
XX AAK06842;
```

```
XX
XX 05-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe SEQ ID NO: 6833.
DE
XX Human; brain expressed exon; gene expression analysis; probe;
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer; ss.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00667.
XX
XX 04-FEB-2000; 2000US-0180312.
XX
XX 26-MAY-2000; 2000US-0207456.
XX
XX 30-JUN-2000; 2000US-0608408.
XX
XX 03-AUG-2000; 2000US-0632366.
XX
XX 21-SEP-2000; 2000US-0234687.
XX
XX 27-SEP-2000; 2000US-0236359.
XX
XX 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
XX Example 4; SEQ ID NO: 6833; 650pp + Sequence listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention.
XX
XX Sequence 484 BP; 161 A; 77 C; 79 G; 167 T; 0 other;
SQ
Query Match 5.9%; Score 160; DB 22; Length 484;
Best Local Similarity 100.0%; Pred. No. 2.6e-68;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 310 AGAATGTTGGAATCCAAATGATGATCACTTCAATGCTTGGCCAAAGCTCCAGTTA 369
DB 458 AGAATGTTGGAATCCAAATGATGATCACTTCAATGCTTGGCCAAAGCTCCAGTTA 399
QY 370 TCATACCTTCTTTTGGATGAGAACGAGTAGAGCTGTATGTTGAGCAAAAGATCACAT 429
DB 398 TCATACCTTCTTTTGGATGAGAACGAGTAGAGCTGTATGTTGAGCAAAAGATCACAT 339
QY 430 ATTTTCATTGACCTGGTTAATATCAAGATTTTCAAAAG 469
DB 338 ATTTTCATTGACCTGGTTAATATCAAGATTTTCAAAAG 299
RESULT 20
ID AAK32561/C
XX AAK32561 standard; DNA; 484 BP.
XX
XX AAK32561;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed single exon probe SEQ ID NO: 7118.
XX
```

KW Human; bone marrow expressed exon; gene expression analysis; probe;
 XX microarray; cancer; leukaemia; lymphoma; myeloma; ss.
 OS Homo sapiens.
 XX WO200157276-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00668.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 XX
 PR 26-MAY-2000; 2000US-0207456.
 XX
 PR 30-JUN-2000; 2000US-0608408.
 XX
 PR 03-AUG-2000; 2000US-0632366.
 XX
 PR 21-SEP-2000; 2000US-0234687.
 XX
 PR 27-SEP-2000; 2000US-0236359.
 XX
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488900/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human bone marrow -
 PS Example 4; SEQ ID NO: 7118; 658bp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
 CC the probes of the invention.
 CC
 SQ Sequence 484 BP; 161 A; 77 C; 79 G; 167 T; 0 other;
 XX
 Query Match 5.9%; Score 160; DB 22; Length 484;
 Best Local Similarity 100.0%; Pred. No. 2.6e-68;
 Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 310 AGAATGTTGGAATCCAAATGTGATCACTTTCAATGCTGGCCAAAGCTCCAGTTA 369
 DB 458 AGAATGTTGGAATCCAAATGTGATCACTTTCAATGCTGGCCAAAGCTCCAGTTA 399
 QY 370 TCATACCTTCCTTTGGATGAGAACGAGTAGGCTGTATGTTGGAGCAAGGATCACAT 429
 DB 398 TCATACCTTCCTTTGGATGAGAACGAGTAGGCTGTATGTTGGAGCAAGGATCACAT 339
 QY 430 ATTTTCATTCGACCTGTTAATATCAAGATTTTCAAAAG 469
 DB 338 ATTTTCATTCGACCTGTTAATATCAAGATTTTCAAAAG 299
 XX
 RESULT 21
 AA138386/c
 ID AA138386 standard; DNA; 484 BP.
 XX
 AC AA138386;
 XX
 DT 17-OCT-2001 (first entry)
 XX
 DE Probe #7072 used to measure gene expression in human placenta sample.
 XX
 KW Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200157272-A2.
 XX

PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00663.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 XX
 PR 26-MAY-2000; 2000US-0207456.
 XX
 PR 30-JUN-2000; 2000US-0608408.
 XX
 PR 03-AUG-2000; 2000US-0632366.
 XX
 PR 21-SEP-2000; 2000US-0234687.
 XX
 PR 27-SEP-2000; 2000US-0236359.
 XX
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-48897/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human placenta -
 PS Claim 25; SEQ ID NO 7072; 654bp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENPs).
 CC The present sequence is one such probe. The probes are useful for
 CC producing a microarray for predicting, measuring and displaying gene
 CC expression in samples derived from human placenta. The probes are useful
 CC for antenatal diagnosis of human genetic disorders.
 CC
 SQ Sequence 484 BP; 161 A; 77 C; 79 G; 167 T; 0 other;
 XX
 Query Match 5.9%; Score 160; DB 22; Length 484;
 Best Local Similarity 100.0%; Pred. No. 2.6e-68;
 Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 310 AGAATGTTGGAATCCAAATGTGATCACTTTCAATGCTGGCCAAAGCTCCAGTTA 369
 DB 458 AGAATGTTGGAATCCAAATGTGATCACTTTCAATGCTGGCCAAAGCTCCAGTTA 399
 QY 370 TCATACCTTCCTTTGGATGAGAACGAGTAGGCTGTATGTTGGAGCAAGGATCACAT 429
 DB 398 TCATACCTTCCTTTGGATGAGAACGAGTAGGCTGTATGTTGGAGCAAGGATCACAT 339
 QY 430 ATTTTCATTCGACCTGTTAATATCAAGATTTTCAAAAG 469
 DB 338 ATTTTCATTCGACCTGTTAATATCAAGATTTTCAAAAG 299
 XX
 RESULT 22
 ABS32272/c
 ID ABS32272 standard; DNA; 484 BP.
 XX
 AC ABS32272;
 XX
 DT 25-FEB-2003 (first entry)
 XX
 DE Human liver single exon probe, SEQ ID NO 7262.
 XX
 KW Human; single exon nucleic acid probe; liver; cirrhosis;
 KW hyperlipoproteinemia; hyperlipidaemia; hypercholesterolaemia;
 KW coronary heart disease; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200157273-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00664.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 XX
 PR 26-MAY-2000; 2000US-0207456.
 XX
 PR 30-JUN-2000; 2000US-0608408.
 XX

PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI, 2001-488898/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analysing gene expression in human adult liver -
 XX
 PS Claim 1; SEQ ID No 7262; 658bp; English.
 XX
 CC The invention relates to a single exon nucleic acid probe (SENP) (1) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult
 CC liver. (1) may be used for predicting, measuring and displaying gene
 CC expression in samples derived from human adult liver. The genes
 CC identified may be involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
 CC is associated with coronary heart disease. ABS25011-ABS51005 represent
 CC human liver single exon nucleic acid probes of the invention.
 CC Note: The sequence information for this patent does not appear in the
 CC printed specification but was obtained in electronic format directly
 CC from WPI at ftp.wipo.int/pub/published_pct_sequences.
 CC
 SQ Sequence 484 BP; 161 A; 77 C; 79 G; 167 T; 0 other;
 XX
 Query Match 5.9%; Score 160; DB 23; Length 484;
 Best Local Similarity 100.0%; Pred. No. 2.6e-68;
 Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 310 AGAATGTTGGATCAACAATGATGATCACTTCAATGGCTTGCCCAACAGCTCCAGTTA 369
 DB 458 AGAATGTTGGATCAACAATGATGATCACTTCAATGGCTTGCCCAACAGCTCCAGTTA 399
 QY 370 TCATACCTTCTTTGGATGAGAAAGGAGTGGCTGATGTTGGAGCAAGATCAT 429
 DB 398 TCATACCTTCTTTGGATGAGAAAGGAGTGGCTGATGTTGGAGCAAGATCAT 339
 QY 430 ATTTTCATTGACCGCTTAATATCAAGATTTTCAAG 469
 DB 338 ATTTTCATTGACCGCTTAATATCAAGATTTTCAAG 299
 XX
 RESULT 23
 ID ABS07351/C
 XX ABS07351 standard; DNA; 484 BP.
 AC ABS07351;
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human genome-derived single exon probe from lung SEQ ID No 7342.
 XX
 KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberos scleriosis; Gaucher's disease; Niemann-Pick disease;
 KW Heremansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenen syndrome;
 KW primary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.

XX
 PD 15-NOV-2001.
 XX
 PE 30-JAN-2001; 2001WO-US00665.
 XX
 XX 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI, 2002-114183/15.
 XX
 PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 XX
 PS Claim 1; SEQ ID No 7342; 634bp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarray having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberos scleriosis, Gaucher's disease,
 CC Niemann-Pick disease, Heremansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagenen syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WPI at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 484 BP; 161 A; 77 C; 79 G; 167 T; 0 other;
 XX
 Query Match 5.9%; Score 160; DB 24; Length 484;
 Best Local Similarity 100.0%; Pred. No. 2.6e-68;
 Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 310 AGAATGTTGGATCAACAATGATGATCACTTCAATGGCTTGCCCAACAGCTCCAGTTA 369
 [Sequence alignment details follow, showing matches and gaps between the query and database sequences.]

Db 458 AGAATGTTGGAAATCCAAATGATGATCACTTCAATGGCTTGGCCAAAGCTCCAGTTA 399
QY 370 TCATACCTTCTTTGGATGAGAACGAGTAGCTGTATGTTGAGCAAGATCAAT 429
Db 398 TCATACCTTCTTTGGATGAGAACGAGTAGCTGTATGTTGAGCAAGATCAAT 339
QY 430 ATTTTCATTCGACCTGGTTAATATCAAGATTTTCAAAAG 469
Db 338 ATTTTCATTCGACCTGGTTAATATCAAGATTTTCAAAAG 299

RESULT 24
AAK07722
ID AAK07722 standard; DNA; 446 BP.
XX
XX AAK07722;
AC
XX 05-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe SEQ ID NO: 7713.
DE
XX Human; brain expressed exon; gene expression analysis; probe;
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.
XX
XX Homo sapiens.
OS
XX
XX MO200157275-A2.
PN
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00667.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR
XX 26-MAY-2000; 2000US-0207456.
PR
XX 30-JUN-2000; 2000US-0608408.
PR
XX 03-AUG-2000; 2000US-0632366.
PR
XX 21-SEP-2000; 2000US-0234687.
PR
XX 27-SEP-2000; 2000US-0236359.
PR
XX 04-OCT-2000; 2000GB-0024263.
XX
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX
XX WPI; 2001-483446/52.
DR
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
XX Example 4; SEQ ID NO: 7713; 650bp + Sequence Listing; English.
PS
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.
XX
XX
XX Sequence 446 BP; 142 A; 88 C; 85 G; 131 T; 0 other;
SQ

Query Match 5.3%; Score 144; DB 22; Length 446;
Best Local Similarity 100.0%; Pred. No. 2e-60;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1916 GATAATCACCATGGCCACAGCCCTGAAGAGAGATCATCTATGTTAGAGATAGTAC 1975
Db 277 GATAATCACCATGGCCACAGCCCTGAAGAGAGATCATCTATGTTAGAGATAGTAC 336
QY 1976 ACATTTTGGATGATGAGTCCGAAGTGGCAGAGAGCGCTGTCTATTGGCAATTCCAGAG 2035
Db 337 ACATTTTGGATGATGAGTCCGAAGTGGCAGAGAGCGCTGTCTATTGGCAATTCCAGAG 396

QY 2036 CGAATGAAGAGCGAAAGAGAG 2059
Db 397 CGAATGAAGAGCGAAAGAGAG 420

RESULT 25
ABS33341
ID ABS33341 standard; DNA; 446 BP.
XX
XX ABS33341;
AC
XX 25-FEB-2003 (first entry)
XX
XX Human liver single exon probe, SEQ ID No 8331.
DE
XX Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
XX Homo sapiens.
OS
XX
XX MO200157273-A2.
PN
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US00664.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR
XX 26-MAY-2000; 2000US-0207456.
PR
XX 30-JUN-2000; 2000US-0608408.
PR
XX 03-AUG-2000; 2000US-0632366.
PR
XX 21-SEP-2000; 2000US-0234687.
PR
XX 27-SEP-2000; 2000US-0236359.
PR
XX 04-OCT-2000; 2000GB-0024263.
XX
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX
XX WPI; 2001-488898/53.
DR
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analysing gene expression in human adult liver -
XX
XX
XX Claim 1; SEQ ID No 8331; 658bp; English.
PS
XX The invention relates to a single exon nucleic acid probe (SENP) (1) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 11109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult
CC liver. (1) may be used for predicting, measuring and displaying gene
CC expression in samples derived from human adult liver. The genes
CC identified may be involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinemia, hyperlipidaemia and hypercholesterolaemia which
CC is associated with coronary heart disease. ABS2011-ABS51005 represent
CC human liver single exon nucleic acid probes of the invention.
CC Note: The sequence information for this patent does not appear in the
CC printed specification but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX
XX Sequence 446 BP; 142 A; 88 C; 85 G; 131 T; 0 other;
SQ

Query Match 5.3%; Score 144; DB 23; Length 446;
Best Local Similarity 100.0%; Pred. No. 2e-60;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1916 GATAATCACCATGGCCACAGCCCTGAAGAGAGATCATCTATGTTAGAGATAGTAC 1975
Db 277 GATAATCACCATGGCCACAGCCCTGAAGAGAGATCATCTATGTTAGAGATAGTAC 336
QY 1976 ACATTTTGGATGATGAGTCCGAAGTGGCAGAGAGCGCTGTCTATTGGCAATTCCAGAG 2035

Db 337 ACATTTTGGATGATGATCCGAAATGCGAGAGGCGTGTCTATTGCAATTCAGAG 396
 OY 2036 CGAATGAAGAGCGAAAAGAGAG 2059
 Db 397 CGAATGAAGAGCGAAAAGAGAG 420

RESULT 26
 ABS08426 ID ABS08426 standard; DNA; 446 BP.
 AC ABS08426;
 DT 19-AUG-2002 (first entry)
 DE Human genome-derived single exon probe from lung SEQ ID No 8417.
 XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Kartagener syndrome;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX Homo sapiens.
 OS WO200186003-A2.
 PN 15-NOV-2001.
 PD 30-JAN-2001; 2001WO-US00665.
 PE 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2002-114183/15.
 DR Spatially-addressable set of single exon nucleic acid probes, used to
 XX measure gene expression in human lung samples -
 PT Claim 1; SEQ ID No 8417; 634bp; English.
 PS The invention relates to a spatially-addressable set of single exon
 XX nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method

CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Kartagener syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 446 BP; 142 A; 88 C; 85 G; 131 T; 0 other;
 SQ

Query Match 5.3%; Score 144; DB 24; Length 446;
 Best Local Similarity 100.0%; Pred. No. 2e-60;
 Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1916 GATATACCATGCGCAGAGCCTGAAAGAGAAATCATCTATGCTAGAGAAATGATAGC 1975
 Db 277 GATATACCATGCGCAGAGCCTGAAAGAGAAATCATCTATGCTAGAGAAATGATAGC 336
 OY 1976 ACATTTTGAATGCAATCCGAAAGCGCTGTCTATTGGCAATTCAGAGG 2035
 Db 337 ACATTTTGAATGCAATCCGAAAGCGCTGTCTATTGGCAATTCAGAGG 396
 OY 2036 CGAATGAAGAGCGAAAAGAGAG 2059
 Db 397 CGAATGAAGAGCGAAAAGAGAG 420

RESULT 27
 ID AAK20386
 AC AAK20386 standard; DNA; 123 BP.
 AC AAK20386;
 DT 05-NOV-2001 (first entry)
 DE Human brain expressed single exon probe SEQ ID NO: 20377.
 XX Human; brain expressed exon; gene expression analysis; probe;
 KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
 KW epilepsy; cancer; ss.
 XX Homo sapiens.
 OS WO200157275-A2.
 PN 09-AUG-2001.
 PD 30-JAN-2001; 2001WO-US00667.
 PE 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.

```
XX Penn SG, Hanzel DK, Chen W, Rank DR,
XX
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
XX Example 4; SEQ ID NO: 20377; 650bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expressions in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention.
XX
XX Sequence 123 BP; 42 A; 19 C; 37 G; 25 T; 0 other;
XX
XX
XX Query Match 4.5%; Score 123; DB 22; Length 123;
XX Best Local Similarity 100.0%; Pred. No. 4.9e-50;
XX Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1937 CCTGAAGAGAGATCTATGCTGTAGAGATAGTACACATTTTGGATGCGCTCG 1996
XX 1 CCTGAAGAGAGATCTATGCTGTAGAGATAGTACACATTTTGGATGCGCTCG 60
XX
XX DB 1997 AAGTGCAGAGAGCGCTGTCTATTGGCAATCCAGAGCGCAATGAAGCGAAAGAA 2056
XX 61 AAGTGCAGAGAGCGCTGTCTATTGGCAATCCAGAGCGCAATGAAGCGAAAGAA 120
XX
XX QY 2057 GAG 2059
XX |||
XX DB 121 GAG 123
XX
XX RESULT 28
XX ABS46246
XX ID ABS46246 standard; DNA; 123 BP.
XX
XX AC ABS46246;
XX
XX DT 25-FEB-2003 (first entry)
XX
XX DE Human liver single exon probe, SEQ ID No 21236.
XX
XX KW Human; single exon nucleic acid probe; liver; cirrhosis;
XX hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
XX coronary heart disease; ss.
XX
XX OS Homo sapiens.
XX
XX PN WO200157273-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US00664.
XX
XX PR 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488898/53.
XX
```

```
PT Human genome-derived single exon nucleic acid probes useful for
PT analysing gene expression in human adult liver -
XX
XX Claim 4; SEQ ID No 21236; 658bp; English.
XX
XX CC The invention relates to a single exon nucleic acid probe (SENP) (1) for
XX CC measuring human gene expression in a sample derived from human adult
XX CC liver, comprising one of 13109 defined nucleotide sequences given in the
XX CC specification (or complements/ fragments). The probe hybridises at high
XX CC stringency to a nucleic acid molecule expressed in the human adult
XX CC liver. (1) may be used for predicting, measuring and displaying gene
XX CC expression in samples derived from human adult liver. The genes
XX CC identified may be involved in genetic liver diseases such as cirrhosis,
XX CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
XX CC is associated with coronary heart disease. ABS25011-ABS51005 represent
XX CC human liver single exon nucleic acid probes of the invention.
XX CC Note: The sequence information for this patent does not appear in the
XX CC printed specification but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.
XX
XX
XX SQ Sequence 123 BP; 42 A; 19 C; 37 G; 25 T; 0 other;
XX
XX
XX Query Match 4.5%; Score 123; DB 23; Length 123;
XX Best Local Similarity 100.0%; Pred. No. 4.9e-50;
XX Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1937 CCTGAAGAGAGATCTATGCTGTAGAGATAGTACACATTTTGGATGCGCTCG 1996
XX 1 CCTGAAGAGAGATCTATGCTGTAGAGATAGTACACATTTTGGATGCGCTCG 60
XX
XX DB 1997 AAGTGCAGAGAGCGCTGTCTATTGGCAATCCAGAGCGCAATGAAGCGAAAGAA 2056
XX 61 AAGTGCAGAGAGCGCTGTCTATTGGCAATCCAGAGCGCAATGAAGCGAAAGAA 120
XX
XX QY 2057 GAG 2059
XX |||
XX DB 121 GAG 123
XX
XX RESULT 29
XX ABS20840
XX ID ABS20840 standard; DNA; 123 BP.
XX
XX AC ABS20840;
XX
XX DT 19-AUG-2002 (first entry)
XX
XX DE Human genome-derived single exon ORF from lung SEQ ID No 20831.
XX
XX KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease; open reading frame; ORF.
XX
XX OS Homo sapiens.
XX
XX PN WO200186003-A2.
XX
XX PD 15-NOV-2001.
XX
XX PF 30-JAN-2001; 2001WO-US00665.
XX
XX PR 04-FEB-2000; 2000US-180312P.
XX PR 26-MAY-2000; 2000US-207456P.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-234687P.
XX PR 27-SEP-2000; 2000US-236359P.
XX
```

PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI
 XX WPI; 2002-114183/15.
 DR

PT Spatially-addressable set of single exon nucleic acid probes, used to
 measure gene expression in human lung samples -

XX Claim 4; SEQ ID NO 20831; 634bp; English.

XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridize at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridization of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridization to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemorrhoidosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagazer syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe open reading frame of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 XX Sequence 123 BP; 42 A; 19 C; 37 G; 25 T; 0 other;

Query Match 4.5%; Score 123; DB 24; Length 123;
 Best Local Similarity 100.0%; Pred. No. 4.9e-50;
 Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1937 CCTGAAGAGATCATCTATGTGTAGAGATAGACATTTTGAATGACAGTCG 1996
 Db 1 CCTGAAGAGATCATCTATGTGTAGAGATAGACATTTTGAATGACAGTCG 60

QY 1997 AAGTCGAGAGAGCGCTGTATTTGCAATTTCCAGAGCCGAATGAAGCGAAGAA 2056
 Db 61 AAGTCGAGAGAGCGCTGTATTTGCAATTTCCAGAGCCGAATGAAGCGAAGAA 120

QY 2057 GAG 2059
 Db 121 GAG 123

RESULT 30
 ABA58557/C
 ID ABA58557 standard; DNA; 456 BP.

XX ABA58557;

XX 01-FEB-2002 (first entry)

DE Human foetal liver single exon nucleic acid probe #6862.

KM Human, foetal liver; gene expression; single exon nucleic acid probe; 88.

XX Homo sapiens.

PN WO200157277-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2001-483447/52.

DR Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human fetal liver -

XX Claim 1; SEQ ID NO 6862; 639bp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human foetal

CC liver. The single exon nucleic acid probes may be used for predicting,

CC measuring and displaying gene expression in samples derived from human

CC foetal liver. The present sequence is a single exon nucleic acid

CC probe of the invention.

CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX
 XX Sequence 456 BP; 150 A; 87 C; 87 G; 132 T; 0 other;

Query Match 4.5%; Score 122; DB 22; Length 456;
 Best Local Similarity 100.0%; Pred. No. 1.5e-49;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 532 GAAGAATGTGCTATTTCAAGTACTTAAGCATATATATGACATCTGTCGC 591
 Db 407 GAAGAATGTGCTATTTCAAGTACTTAAGCATATATATGACATCTGTCGC 348

QY 592 CTGTGAACGGGGCTTTTCAATTCATCCAAATTTGCACTTACATTTGAATGACATCTGA 651
 Db 347 CTGTGAACGGGGCTTTTCAATTCATCCAAATTTGCACTTACATTTGAATGACATCTGA 288

QY 652 GG 653
 Db 287 GG 286

RESULT 31
 ABA27590/C
 ID ABA27590 standard; DNA; 456 BP.
 XX
 XX ABA27590;

XX 23-JAN-2002 (first entry)
DT
XX
XX Probe #6056 for gene expression analysis in human heart cell sample.
DE
XX Human; gene expression; heart; microarray; vascular system; probe;
XX cardiovascular disease; hypertension; cardiac arrhythmia;
XX congenital heart disease; ss.
XX
XX Homo sapiens.
XX
XX WO200157274-A2.
XX
XX
XX PD 09-AUG-2001.
XX
XX
XX PF 30-JAN-2001; 2001WO-US00666.
XX
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI WPI; 2001-488899/53.
XX DR
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX PT hearts -
XX
XX PS Claim 1; SEQ ID NO 6056; 530bp; English.
XX
XX CC The present invention relates to single exon nucleic acid probes for
XX CC measuring human gene expression in a sample derived from human heart. The
XX CC present sequence is one such probe. The probes may be used for
XX CC predicting, measuring and displaying gene expression in samples derived
XX CC from the human heart via microarrays. By measuring gene expression, the
XX CC probes are useful for predicting, diagnosing, grading, staging,
XX CC monitoring and prognosing diseases of the human heart and vascular system
XX CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
XX CC congenital heart disease.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pcc_sequences.
XX
XX SQ Sequence 456 BP; 150 A; 87 C; 87 G; 132 T; 0 other;
XX
XX Query Match 4.5%; Score 122; DB 22; Length 456;
XX Best Local Similarity 100.0%; Pred. No. 1.5e-49;
XX Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 532 GAAGAATGTGCTAATTCATCAAGTACTTAAGGATATATACAGCTCCTTGTACGC 591
XX DB 407 GAAGAATGTGCTAATTCATCAAGTACTTAAGGATATATACAGCTCCTTGTACGC 348
XX
XX QY 592 CTGTGGAACGGGGGCTTTTCATCAATTTGCACCTACATTAATTTGACATCATCTCTGA 651
XX DB 347 CTGTGGAACGGGGGCTTTTCATCAATTTGCACCTACATTAATTTGACATCATCTCTGA 288
XX
XX QY 652 GG 653
XX DB 287 GG 286

RESULT 32
AAK06676/C
ID . AAK06676 standard; DNA; 456 BP.
XX
XX AAK06676;
AC

XX 05-NOV-2001 (first entry)
DT
XX
XX Human brain expressed single exon probe SEQ ID NO: 6667.
DE
XX Human; brain expressed exon; gene expression analysis; probe;
XX microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
XX epilepsy; cancer; ss.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX
XX PD 09-AUG-2001.
XX
XX
XX PF 30-JAN-2001; 2001WO-US00667.
XX
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI WPI; 2001-483446/52.
XX DR
XX PT Single exon nucleic acid probes for analyzing gene expression in human
XX PT brains -
XX
XX PS Example 4; SEQ ID NO: 6667; 650bp + Sequence listing; English.
XX
XX CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC brain. They can be used to measure gene expression in brain cell samples,
XX CC which may enable the diagnosis and improved treatment of nervous system
XX CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX CC epilepsy and cancer. The present sequence is one of the probes of the
XX CC invention.
XX
XX SQ Sequence 456 BP; 150 A; 87 C; 87 G; 132 T; 0 other;
XX
XX Query Match 4.5%; Score 122; DB 22; Length 456;
XX Best Local Similarity 100.0%; Pred. No. 1.5e-49;
XX Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 532 GAAGAATGTGCTAATTCATCAAGTACTTAAGGATATATACAGCTCCTTGTACGC 591
XX DB 407 GAAGAATGTGCTAATTCATCAAGTACTTAAGGATATATACAGCTCCTTGTACGC 348
XX
XX QY 592 CTGTGGAACGGGGGCTTTTCATCAATTTGCACCTACATTAATTTGACATCATCTCTGA 651
XX DB 347 CTGTGGAACGGGGGCTTTTCATCAATTTGCACCTACATTAATTTGACATCATCTCTGA 288
XX
XX QY 652 GG 653
XX DB 287 GG 286

RESULT 33
AAK32372/C
ID AAK32372 standard; DNA; 456 BP.
XX
XX AAK32372;
XX
XX AC 06-NOV-2001 (first entry)
XX DT Human bone marrow expressed single exon probe SEQ ID NO: 6929.
XX DE
XX

KW Human; bone marrow expressed exon; gene expression analysis; probe;
 KW microarray; cancer; leukemia; lymphoma; myeloma; ss.
 OS Homo sapiens.
 XX MO200157276-A2.
 PN 09-AUG-2001.
 PD 30-JAN-2001; 2001WO-US00668.
 PF 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI: 2001-488900/53.
 DR WPI: 2001-488900/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human bone marrow -
 CC Example 4; SEQ ID NO: 6929; 658bp + Sequence listing; English.
 PS The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukemia and myeloma. The present sequence is one of
 CC the probes of the invention.
 XX
 SQ Sequence 456 BP; 150 A; 87 C; 87 G; 132 T; 0 other;
 QY Query Match 4.5%; Score 122; DB 22; Length 456;
 Db Best Local Similarity 100.0%; Pred. No. 1.5e-49; Indels 0; Gaps 0;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 532 GAAAGAAATGCTAATTCATCAAGTACTTAAGGCATATATGAGCTCACTTGTACGC 591
 Db 407 GAAAGAAATGCTAATTCATCAAGTACTTAAGGCATATATGAGCTCACTTGTACGC 348
 QY 592 CTGTGGAACGGGGCTTTTCATCCATTGCACTTGAATTGACATCATCTCTGA 651
 Db 347 CTGTGGAACGGGGCTTTTCATCCATTGCACTTGAATTGACATCATCTCTGA 288
 QY 652 GG 653
 Db 287 GG 286
 DE RESULT 34
 ID AA138218/c
 AC AA138218;
 DT 17-OCT-2001 (first entry)
 DE Probe #6904 used to measure gene expression in human placenta sample.
 KW Probe; microarray; human; placenta; antenatal diagnosis;
 KW genetic disorder; ss.
 OS Homo sapiens.
 XX MO200157272-A2.
 PN

PD 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US00663.
 PF 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI: 2001-488897/53.
 DR WPI: 2001-488897/53.
 XX Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human placenta -
 CC Claim 25; SEQ ID No 6904; 654bp; English.
 PS The present invention relates to single exon nucleic acid probes (SENP).
 CC The present sequence is one such probe. The probes are useful for
 CC producing a microarray for predicting, measuring and displaying gene
 CC expression in samples derived from human placenta. The probes are useful
 CC for antenatal diagnosis of human genetic disorders.
 XX
 SQ Sequence 456 BP; 150 A; 87 C; 87 G; 132 T; 0 other;
 QY Query Match 4.5%; Score 122; DB 22; Length 456;
 Db Best Local Similarity 100.0%; Pred. No. 1.5e-49; Indels 0; Gaps 0;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 532 GAAAGAAATGCTAATTCATCAAGTACTTAAGGCATATATGAGCTCACTTGTACGC 591
 Db 407 GAAAGAAATGCTAATTCATCAAGTACTTAAGGCATATATGAGCTCACTTGTACGC 348
 QY 592 CTGTGGAACGGGGCTTTTCATCCATTGCACTTGAATTGACATCATCTCTGA 651
 Db 347 CTGTGGAACGGGGCTTTTCATCCATTGCACTTGAATTGACATCATCTCTGA 288
 QY 652 GG 653
 Db 287 GG 286
 DE RESULT 35
 ID ABS32080/c
 AC ABS32080;
 DT 25-FEB-2003 (first entry)
 DE Human liver single exon probe, SEQ ID No 7070.
 KW Human; single exon nucleic acid probe; liver; cirrhosis;
 KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
 KW coronary heart disease; ss.
 OS Homo sapiens.
 XX MO200157273-A2.
 PN 09-AUG-2001.
 PD 30-JAN-2001; 2001WO-US00664.
 PF 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PN

PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;
 PA WPI; 2001-48898/53.

PT Human genome-derived single exon nucleic acid probes useful for
 PT analysing gene expression in human adult liver -

PS Claim 1; SEQ ID No 7070; 658bp; English.

CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult
 CC liver. (II) may be used for predicting, measuring and displaying gene
 CC expression in samples derived from human adult liver. The genes
 CC identified may be involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
 CC is associated with coronary heart disease. ABS25011-ABS51005 represent
 CC human liver single exon nucleic acid probes of the invention.
 CC Note: The sequence information for this patent does not appear in the
 CC printed specification but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 456 BP; 150 A; 87 C; 87 G; 132 T; 0 other;

Query Match 4.5%; Score 122; DB 23; Length 456;
 Best Local Similarity 100.0%; Pred. No. 1.5e-49;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 532 GAAGATGCTAATTTTCATCAGTACTTAAGCATTAATCAAGTCACTTGACGC 591

DB 407 GAAAGATGCTAATTTTCATCAGTACTTAAGCATTAATCAAGTCACTTGACGC 348

QY 552 CTGTGACGCGGGCTTTTCATCATTGACACTTGAATTGACATCATCTTGA 651

DB 347 CTGTGACGCGGGCTTTTCATCATTGACACTTGAATTGACATCATCTTGA 288

QY 652 GG 653

DB 287 GG 286

RESULT 36
 ABS07155/C
 ID ABS07155 standard; DNA; 456 BP.

AC ABS07155;

DT 19-AUG-2002 (first entry)

DE Human genome-derived single exon probe from lung SEQ ID No 7146.

XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hereditary-Pudlak syndrome; sarcoidosis; pulmonary haemorrhoidosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.

OS Homo sapiens.

XX WO200186003-A2.

XX 15-NOV-2001.
 PD 30-JAN-2001; 2001WO-US00665.

PR 04-FEB-2000; 2000US-180112P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.

(MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;
 PA WPI; 2002-114183/15.

PT Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -

PS Claim 1; SEQ ID No 7146; 634bp; English.

CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hereditary-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemorrhoidosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karsenger syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 456 BP; 150 A; 87 C; 87 G; 132 T; 0 other;

Query Match 4.5%; Score 122; DB 24; Length 456;
 Best Local Similarity 100.0%; Pred. No. 1.5e-49;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 532 GAAGATGCTAATTTTCATCAGTACTTAAGCATTAATCAAGTCACTTGACGC 591

|||||

Db 407 GAAAGATGTGCTAATTCATCAAGTACTTAAGGCATTAATTCAGACTCACTTGACGC 348
 QY 592 CTGTGGAACGGGGGCTTTTCATCAATTTGACACTTGAATTTGACATCATCTGCA 651
 Db 347 CTGTGGAACGGGGGCTTTTCATCAATTTGACACTTGAATTTGACATCATCTGCA 288
 QY 652 GG 653
 Db 287 GG 286

RESULT 37

ABA71291/c
 ID ABA71291 standard; DNA; 101 BP.

AC ABA71291;
 XX

DT 01-FEB-2002 (first entry)
 XX

DE Human foetal liver single exon nucleic acid probe #19596.
 XX

KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
 XX

OS Homo sapiens.
 XX

PN WO200157277-A2.
 XX

PD 09-AUG-2001.
 XX

PF 30-JAN-2001; 2001WO-US00669.
 XX

PR 04-FEB-2000; 2000US-0180312.
 XX

PR 26-MAY-2000; 2000US-0207456.
 XX

PR 30-JUN-2000; 2000US-0608408.
 XX

PR 03-AUG-2000; 2000US-0632366.
 XX

PR 21-SEP-2000; 2000US-0234687.
 XX

PR 27-SEP-2000; 2000US-0236359.
 XX

PR 04-OCT-2000; 2000GB-0024263.
 XX

PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX

DR WPI; 2001-483447/52.
 XX

PT Human genome-derived single exon nucleic acid probes useful for
 XX

PS analyzing gene expression in human fetal liver.
 XX

PS Claim 4; SEQ ID NO 19596; 639pp + sequence listing; English.
 XX

CC The invention relates to a single exon nucleic acid probe for
 XX

CC measuring human gene expression in a sample derived from human foetal
 XX

CC liver. The single exon nucleic acid probes may be used for predicting,
 XX

CC measuring and displaying gene expression in samples derived from human,
 XX

CC foetal liver. The present sequence is a single exon nucleic acid
 XX

CC probe of the invention.
 XX

CC Note: The sequence data for this patent did not form part of the
 XX

CC printed specification, but was obtained in electronic format directly
 XX

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX

CC Sequence 101 BP; 30 A; 20 C; 21 G; 30 T; 0 other;
 XX

SO Query Match 3.7%; Score 101; DB 22; Length 101;
 XX

Best Local Similarity 100.0%; Pred. No. 3.6e-39;
 XX

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX

QY 533 AAAAGATGTGCTAATTCATCAAGTACTTAAGGCATTAATTCAGACTCACTTGACGC 592
 XX

Db 101 AAAAGATGTGCTAATTCATCAAGTACTTAAGGCATTAATTCAGACTCACTTGACGC 42
 XX

QY 593 TGTGGAACGGGGGCTTTTCATCAATTTGACACTTGAATTTGACATCATCTGCA 633
 XX

Db 41 TGTGGAACGGGGGCTTTTCATCAATTTGACACTTGAATTTGACATCATCTGCA 1
 XX

RESULT 38
 ABA37573/c
 ID ABA37573 standard; DNA; 101 BP.

AC ABA37573;
 XX

DT 23-JAN-2002 (first entry)
 XX

DE Probe #16039 for gene expression analysis in human heart cell sample.
 XX

KW Human; gene expression; heart; microarray; vascular system; probe;
 XX

KW cardiovascular disease; hypertension; cardiac arrhythmia;
 XX

OS Homo sapiens.
 XX

PN WO200157274-A2.
 XX

PD 09-AUG-2001.
 XX

PF 30-JAN-2001; 2001WO-US00669.
 XX

PR 04-FEB-2000; 2000US-0180312.
 XX

PR 26-MAY-2000; 2000US-0207456.
 XX

PR 30-JUN-2000; 2000US-0608408.
 XX

PR 03-AUG-2000; 2000US-0632366.
 XX

PR 21-SEP-2000; 2000US-0234687.
 XX

PR 27-SEP-2000; 2000US-0236359.
 XX

PR 04-OCT-2000; 2000GB-0024263.
 XX

PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX

DR WPI; 2001-488899/53.
 XX

PT Single exon nucleic acid probes for analyzing gene expression in human
 XX

PS hearts -
 XX

PS Claim 4; SEQ ID NO 16039; 530pp; English.
 XX

CC The present invention relates to single exon nucleic acid probes for
 XX

CC measuring human gene expression in a sample derived from human heart. The
 XX

CC present sequence is one such probe. The probes may be used for
 XX

CC predicting, measuring and displaying gene expression in samples derived
 XX

CC from the human heart via microarrays. By measuring gene expression, the
 XX

CC probes are useful for predicting, diagnosing, grading, staging,
 XX

CC monitoring and prognosing diseases of the human heart and vascular system
 XX

CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
 XX

CC congenital heart disease.
 XX

CC Note: The sequence data for this patent did not form part of the printed
 XX

CC specification, but was obtained in electronic format directly from WIPO
 XX

CC at ftp.wipo.int/pub/published_pct_sequences.
 XX

CC Sequence 101 BP; 30 A; 20 C; 21 G; 30 T; 0 other;
 XX

SO Query Match 3.7%; Score 101; DB 22; Length 101;
 XX

Best Local Similarity 100.0%; Pred. No. 3.6e-39;
 XX

Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX

QY 533 AAAAGATGTGCTAATTCATCAAGTACTTAAGGCATTAATTCAGACTCACTTGACGC 592
 XX

Db 101 AAAAGATGTGCTAATTCATCAAGTACTTAAGGCATTAATTCAGACTCACTTGACGC 42
 XX

QY 593 TGTGGAACGGGGGCTTTTCATCAATTTGACACTTGAATTTGACATCATCTGCA 633
 XX

Db 41 TGTGGAACGGGGGCTTTTCATCAATTTGACACTTGAATTTGACATCATCTGCA 1
 XX

RESULT 39
 AAK19593/c

```
ID AAK19593 standard; DNA; 101 BP.
XX
AC AAK19593;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 19584.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 19584; 650bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention.
XX
SQ Sequence 101 BP; 30 A; 20 C; 21 G; 30 T; 0 other;
XX
Query Match 3.7%; Score 101; DB 22; Length 101;
Best Local Similarity 100.0%; Pred. No. 3.6e-39;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 533 AAGAGATGCTTAATTTTCATCAAGGTAAGGCTAATATCAAGCTCACTTGTACGCC 592
DB 101 AAGAGATGCTTAATTTTCATCAAGGTAAGGCTAATATCAAGCTCACTTGTACGCC 42
QY 593 TGTGGAACGGGGGCTTTTCATCCAAATTGACCTACATTGA 633
DB 41 TGTGGAACGGGGGCTTTTCATCCAAATTGACCTACATTGA 1
XX
RESULT 40
ID AAK45596/C
XX AAK45596 standard; DNA; 101 BP.
XX
AC AAK45596;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 20153.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
```

```
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 20153; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is one of
CC the probes of the invention.
XX
SQ Sequence 101 BP; 30 A; 20 C; 21 G; 30 T; 0 other;
XX
Query Match 3.7%; Score 101; DB 22; Length 101;
Best Local Similarity 100.0%; Pred. No. 3.6e-39;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 533 AAGAGATGCTTAATTTTCATCAAGGTAAGGCTAATATCAAGCTCACTTGTACGCC 592
DB 101 AAGAGATGCTTAATTTTCATCAAGGTAAGGCTAATATCAAGCTCACTTGTACGCC 42
QY 593 TGTGGAACGGGGGCTTTTCATCCAAATTGACCTACATTGA 633
DB 41 TGTGGAACGGGGGCTTTTCATCCAAATTGACCTACATTGA 1
XX
RESULT 41
ID AA151533/C
XX AA151533 standard; DNA; 101 BP.
XX
AC AA151533;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #20219 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00663.
XX
PR 04-FEB-2000; 2000US-0180312.
```

26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-063366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488897/53.
XX
XX
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human placenta -
XX
XX Claim 25; SEQ ID No 20219; 654bp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SENP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders.
XX
XX Sequence 101 BP; 30 A; 20 C; 21 G; 30 T; 0 other;
SQ
Query Match 3.7%; Score 101; DB 22; Length 101;
Best Local Similarity 100.0%; Pred. No. 3.6e-39;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 533 AAAGATGCTAATTTTCATCAAGGTAAGGATTAATCAAGACTCACTGTACGCC 592
Db 101 AAAGATGCTAATTTTCATCAAGGTAAGGATTAATCAAGACTCACTGTACGCC 42
Qy 593 TGTGAACGGGGGCTTTTCATCAATTTGCACTACCTACATTA 633
Db 41 TGTGAACGGGGGCTTTTCATCAATTTGCACTACCTACATTA 1
RESULT 42
ABS45286/C
ID ABS45286 standard; DNA; 101 BP.
XX
XX ABS45286;
XX
XX 25-FEB-2003 (first entry)
XX
XX Human liver single exon probe, SEQ ID No 20276.
DE Human; single exon nucleic acid probe; liver; cirrhosis;
XX hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
XX Homo sapiens.
OS
XX WO200157273-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00664.
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-063366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX

WPI; 2001-488898/53.
DR
XX Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human adult liver -
XX
XX Claim 4; SEQ ID No 20276; 656bp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult
CC liver. (I) may be used for predicting, measuring and displaying gene
CC expression in samples derived from human adult liver. The genes
CC identified may be involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which
CC is associated with coronary heart disease. ABS25011-ABS51005 represent
CC human liver single exon nucleic acid probes of the invention.
CC Note: The sequence information for this patent does not appear in the
CC printed specification but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 101 BP; 30 A; 20 C; 21 G; 30 T; 0 other;
SQ
Query Match 3.7%; Score 101; DB 23; Length 101;
Best Local Similarity 100.0%; Pred. No. 3.6e-39;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 533 AAAGATGCTAATTTTCATCAAGGTAAGGATTAATCAAGACTCACTGTACGCC 592
Db 101 AAAGATGCTAATTTTCATCAAGGTAAGGATTAATCAAGACTCACTGTACGCC 42
Qy 593 TGTGAACGGGGGCTTTTCATCAATTTGCACTACCTACATTA 633
Db 41 TGTGAACGGGGGCTTTTCATCAATTTGCACTACCTACATTA 1
RESULT 43
ABS19868/C
ID ABS19868 standard; DNA; 101 BP.
XX
XX ABS19868;
XX
XX 19-AUG-2002 (first entry)
XX
XX Human genome-derived single exon probe ORF from lung SEQ ID No 19859.
DE Human; ds; single exon probe; asthma; lung cancer; COPD; IID;
XX chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemorrhoidosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
XX Homo sapiens.
OS
XX WO200186003-A2.
XX
XX 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US00665.
XX
XX 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-063366.
PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX

PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2002-114183/15.
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 PS Claim 4; SEQ ID No 19859; 634bp; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 1287 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridize at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning sequence to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarray having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 1201 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karsagen syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a single exon
 CC probe open reading frame of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPD at
 CC ftp.wipd.int/pub/published_pct_sequences.
 XX
 SQ Sequence 101 BP; 30 A; 20 C; 21 G; 30 T; 0 other;
 XX
 Query Match 3.7%; Score 101; DB 24; Length 101;
 Best Local Similarity 100.0%; Pred. No. 3.6e-39;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 533 AAGAATGCTCAATTTCATCAAGGACTTAAGGCAATATATCAGACTGTTAGCC 592
 Db 101 AAGAATGCTCAATTTCATCAAGGACTTAAGGCAATATATCAGACTGTTAGCC 42
 QY 593 TGTGGAACGGGGCTTTTCATCCAAATTTGCACCTACATGA 633
 Db 41 TGTGGAACGGGGCTTTTCATCCAAATTTGCACCTACATGA 1

DT 29-OCT-2001 (first entry)
 XX
 XX Human maspin cDNA sequence.
 XX
 XX Drug resistance; resistance gene; semaphorin D, B94; mel-14 antigen;
 XX 24p3; proliferin; maspin; cancer; cytoskeletal; gene therapy; ss.
 XX Homo sapiens.
 XX
 PN WO20015455-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 31-JAN-2001; 2001WO-US03161.
 XX
 PR 31-JAN-2000; 2000US-0179191.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 PA (JINS/) JIN S.
 PI Jin S;
 XX
 PS WPI; 2001-488799/53.
 DR
 XX Determining if a compound modulates the drug resistance of a cell,
 PT comprises determining the expression or activity level of a resistance
 PT sequence in a cell in the presence of the test compound -
 XX
 XX Example 1; Fig 5A-B; 79pp; English.
 XX
 CC The invention relates to a method of determining whether a test compound
 CC modulates the drug resistance of a cell that comprises determining the
 CC expression or activity level of resistance genes (e.g. semaphorin D, B94,
 CC mel-14 antigen, 24p3, proliferin or maspin) in a cell in the presence of
 CC the test compound, and comparing its expression or activity level in a
 CC cell without the test compound. The drug resistant sequences are useful
 CC in identifying drug resistant cells, in screening methods directed to the
 CC identification of compounds that can modulate the drug resistance of a
 CC cell type or multiple cell types. An isolated resistance protein can be
 CC used as an immunogen to generate antibodies that bind the resistance
 CC protein. Resistance nucleic acids may be inserted into vectors and used
 CC as gene therapy vectors. An anti-resistance protein antibody may be used
 CC to isolate a resistance protein, or facilitate the purification of
 CC natural resistance protein from cells and of recombinantly produced
 CC resistance protein expressed in host cells. The methods are useful for
 CC treating a subject having a disorder, such as a drug-resistance cancer,
 CC characterized by aberrant resistance sequence expression or activity by
 CC administering to the subject a resistance modulator. The present sequence
 CC represents a human maspin cDNA sequence, whose expression was decreased
 CC in drug resistant EMT6 tumours.
 XX
 SQ Sequence 2136 BP; 658 A; 454 C; 428 G; 596 T; 0 other;
 XX
 Query Match 2.6%; Score 71; DB 22; Length 2136;
 Best Local Similarity 100.0%; Pred. No. 2.2e-24;
 Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AATCTTATTTTATTCAGTGTATACAGCTAGTATGATGATCCAGCGTGGGGTTCG 60
 Db 17 AATCTTATTTTATTCAGTGTATACAGCTAGTATGATGATGATCCAGCGTGGGGTTCG 76
 QY 61 ACCCAGCGTC 71
 Db 77 ACCCAGCGTC 87

RESULT 44
 AAH47053
 ID AAH47053 standard; DNA; 2136 BP.
 XX
 AC AAH47053;
 XX

RESULT 45
 AAC02557
 ID AAC02557 standard; cDNA; 229 BP.
 XX
 AC AAC02557;
 XX
 DT 06-OCT-2000 (first entry)

```

XX XX Human secreted protein 5' EST, SEQ ID NO: 2555.
DE DE
XX XX
XX XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW KW gene therapy; chromosome mapping; ss.
XX XX
OS OS Homo sapiens.
XX XX
XX XX EP1033401-A2.
XX XX
XX XX
XX XX 06-SEP-2000.
XX XX
XX XX 21-FEB-2000; 2000EP-0200610.
XX XX
XX XX 26-FEB-1999; 99US-0122487.
XX XX
XX XX (GSEST ) GENSET.
XX XX
XX XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX XX
XX XX WPI; 2000-500381/45.
XX XX
XX XX P-PSDB; AAG02551.
XX XX
XX XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX XX
XX XX Claim 1, SEQ ID 2555; 71pp + CD-ROM; English.
XX XX
XX XX The present sequence is one of a large number of 5' ESTs derived from
CC CC mRNAs encoding secreted proteins. An ORF has been identified within the
CC CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC CC derived from 30 different tissues. EST sequences usually correspond
CC CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC CC well suited for isolating cDNA sequences derived from the 5' ends of
CC CC mRNAs and even in those cases where longer cDNA sequences have been
CC CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC CC gene therapy and chromosome mapping procedures. They are used to obtain
CC CC upstream regulatory sequences and to design expression and secretion
XX XX vectors.
XX XX
XX XX Sequence 229 BP; 63 A; 47 C; 53 G; 62 T; 4 other;
SQ SQ
XX XX
XX XX Query Match 2.3%; Score 63; DB 21; Length 229;
XX XX Best Local Similarity 100.0%; Pred. No. 2.1e-20;
XX XX Matches 63; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX XX
OY 174 GCAAGGAGACCTACGCTGACAGATGGGCTGTTAAGTATGATGCTTTCT 233
DB 123 GCAAGGAGACCTACGCTGACAGATGGGCTGTTAAGTATGATGCTTTCT 182
XX XX
OY 234 GGG 236
DB 183 GGG 185
XX XX
XX XX
XX XX
XX XX 15-JUL-2002 (first entry)
XX XX
XX XX Human spliced transcript detection oligonucleotide SEQ ID NO:15846.
XX XX
XX XX Human; mouse; rat; splice transcript; detection; RNA transcript;
KW KW splice variant; transcriptome; oligonucleotide library; ss.
XX XX
OS OS Homo sapiens.

```

```

XX XX WO200210449-A2.
XX XX
XX XX
XX XX 07-FEB-2002.
XX XX
XX XX
XX XX 20-JUL-2001; 2001WO-IB01903.
XX XX
XX XX 28-JUL-2000; 2000US-221607P.
XX XX
XX XX 02-MAY-2001; 2001US-287724P.
XX XX
XX XX (COMP-) COMPUGEN INC.
XX XX
XX XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX XX
XX XX WPI; 2002-257383/30.
XX XX
XX XX
XX XX New oligonucleotide libraries comprising oligonucleotides which
PT PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT PT a genome, useful for detecting tissue-, pathology-, and
XX XX developmental-specific genes -
XX XX
XX XX Example 1; SEQ ID 15846; 47pp; English.
XX XX
XX XX The present invention describes oligonucleotide libraries for detecting
CC CC messenger RNAs that populate a (sub-)transcriptome, where the
CC CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC CC transcription units that populate a genome. The library comprises
CC CC several oligonucleotides, each capable of hybridizing selectively to a
CC CC set of messenger RNAs transcribed from a given transcription unit of
CC CC the genome, which encodes one or more messenger RNA splice variants.
CC CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC CC biological sample, in expression profiling studies, in qualitatively or
CC CC quantitatively characterizing the corresponding transcriptome, and in
CC CC detecting RNA transcripts and splice variants of human or animal
CC CC transcripts. The libraries may also be used as specialised mini
CC CC libraries to detect transcripts of a sub-transcriptome under a
CC CC particular biological or pathological state, and so allowing the
CC CC detection of tissue- and pathology-specific genes such as those genes
CC CC only expressed in specific tissue under a specific pathological
CC CC condition, to detect developmental specific genes; and to detect RNA
CC CC transcripts and splice variants of a transcriptome of a patient suffering
CC CC from a particular disorder. ABN27253 to ABN59589 represent
CC CC oligonucleotide sequences from rats, humans and mice, which are used in
CC CC the exemplification of the present invention.
CC CC N.B. The sequence data for this patent did not form part of the printed
CC CC specification, but was obtained in electronic format directly from WIRO
CC CC at ftp.wipo.int/pub/published_poc_sequences.
XX XX
XX XX Sequence 60 BP; 23 A; 11 C; 15 G; 11 T; 0 other;
SQ SQ
XX XX
XX XX Query Match 2.2%; Score 60; DB 24; Length 60;
XX XX Best Local Similarity 100.0%; Pred. No. 6.4e-19;
XX XX Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX XX
OY 950 GCAATAGATGAGAAACACTGGAAGGCTACTACCGCTAGAAATAGATATGCAAG 1009
DB 1 GCAATAGATGAGAAACACTGGAAGGCTACTACCGCTAGAAATAGATATGCAAG 60
XX XX
XX XX
XX XX
XX XX
XX XX 09-JAN-2003 (first entry)
XX XX
XX XX Human leukocyte gene expression profiling probe SEQ ID NO 4731.
XX XX
XX XX T7; leukocyte; gene expression profiling; allograft rejection;
KW KW atherosclerosis; congestive heart failure; systemic lupus erythematosus;
KW KW rheumatoid arthritis; osteoarthritis; cytomegalovirus; infection;
KW KW probe; ss.

```

```

XX OS Homo sapiens.
XX XX MO200257414-A2.
XX PN 25-JUL-2002.
XX PD
XX PF 22-OCT-2001; 2001WO-US47856.
XX XX
XX PR 20-OCT-2000; 2000US-241994P.
XX PR 08-JUN-2001; 2001US-296764P.
XX XX
XX PA (BIOC-) BIOCARDIA INC.
XX PI Wohlgenuth J, Fry K, Matcuk G, Altman P, Prentice J, Phillips J,
XX PI Ly N, Woodward R, Queternous T, Johnson F;
XX XX
XX DR WPI; 2002-636525/68.
XX XX
XX PT New system for leukocyte expression profiling, diagnosing a disease, or
XX PT monitoring (the rate of) progression of a disease, e.g. atherosclerosis
XX PT or congestive heart failure, comprises diagnostic oligonucleotides -
XX XX
XX PS Claim 1, Page 479; 2038pp; English.
XX XX
XX CC The invention relates to a system for detecting gene expression, which
XX CC comprises one or two isolated DNA molecules that detect expression of a
XX CC gene, where the gene corresponds to any of 8143 oligonucleotides
XX CC (AB000010-AB208152) each having 50 base pairs (bp). The system is useful
XX CC for leukocyte expression profiling. It is particularly useful for
XX CC diagnosing a disease, monitoring (rate of) progression of a disease,
XX CC predicting therapeutic outcome, determining prognosis for a patient,
XX CC predicting disease complications in an individual or monitoring response
XX CC to treatment in an individual. The diseases include cardiac allograft
XX CC rejection, kidney allograft rejection, liver allograft rejection,
XX CC atherosclerosis, congestive heart failure, systemic lupus erythematosus,
XX CC rheumatoid arthritis, osteoarthritis or cytomegalovirus infection.
XX XX
XX SQ Sequence 50 BP; 19 A; 13 C; 7 G; 11 T; 0 other;
XX XX
XX QY Query Match 1.8%; Score 50; DB 24; Length 50;
XX QY Best Local Similarity 100.0%; Pred. No. 5.6e-14;
XX QY Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX XX
XX DB 2516 GCTGATTACCTCTAGAAACCTCAACAGTAGAACTTGCTAGACAAT 2565
XX |
XX | 1 GCTGATTACCTCTAGAAACCTCAACAGTAGAACTTGCTAGACAAT 50
XX |
XX XX
XX RESULT 48
XX AAD28273/C
XX ID AAD28273 standard; DNA; 4661 BP.
XX XX
XX AC AAD28273;
XX XX
XX DT 22-APR-2002 (first entry)
XX XX
XX DE Alpha-lactalbumin (LA) cc49IL2 vector.
XX XX
XX KM Bovine; alpha-lactalbumin; promoter; pharmaceutical; industrial; human;
XX KM diagnostic; screening; moloney murine leukemia virus; MoMuLV; chimeric;
XX KM ds.
XX XX
XX OS Chimeric - Bos sp.
XX OS Chimeric - Homo sapiens.
XX OS Chimeric - Moloney murine leukemia virus.
XX OS Chimeric - Unidentified.
XX XX
XX FH Key Location/Qualifiers
XX FT 1..2055
XX FT misc_feature /tag= a
XX FT /note= "Bovine/human alpha-lactalbumin 5' flanking
XX FT region"

```

```

FT CDS 2098..4011
FT FT /*tag= b
FT FT /product= "cc49-IL2 coding region"
FT LTR 4068..4661
FT FT /*tag= c
FT FT /note= "MoMuLV 3' LTR"
XX XX
XX PN WO200202738-A2.
XX XX
XX PD 10-JAN-2002.
XX XX
XX PF 29-JUN-2001; 2001WO-US20710.
XX XX
XX PR 03-JUL-2000; 2000US-215925P.
XX PR (GALA-) GALA DESIGN INC.
XX PA
XX PI Bremel RD, Miller LU, Bleck GT, York D;
XX PI WPI; 2002-154737/20.
XX XX
XX DR Host cell for producing a desired protein and for screening compounds
XX DR useful for pharmaceutical, industrial, diagnostic and other purposes,
XX DR comprises multiple integrating vectors having an exogenous gene -
XX XX
XX PS Example 1; Fig 13; 191pp; English.
XX XX
XX CC The invention relates to a host cell comprising a genome having at least
XX CC two integrated integrating vectors. The integrating vectors comprise at
XX CC least one exogenous gene operably linked to a promoter. The host cell
XX CC is useful for producing a desired protein and for comparing protein
XX CC functions. The host cells comprises a reporter gene which is from gene
XX CC fluorescent protein, luciferase, beta-galactosidase and beta-lactamase,
XX CC and the assaying step further comprises detecting a signal from the
XX CC reporter gene. The desired protein includes proteins for pharmaceutical,
XX CC industrial, diagnostic and other purposes. The host cells is useful for
XX CC indirectly detecting the expression of a desired protein, comprising
XX CC providing the host cell transfected with a vector encoding a
XX CC polycistronic sequence comprising a signal protein and a desired protein
XX CC operably linked by an internal ribosome entry site (IRES), and culturing
XX CC the host cell under suitable conditions so that the signal protein and
XX CC the desired protein is produced, where the presence of signal protein
XX CC indicates the presence of desired protein. The present sequence is
XX CC alpha lactalbumin (LA) cc49IL2 vector used in the invention. The vector
XX CC comprises the following elements: bovine/human alpha-lactalbumin hybrid
XX CC promoter, cc49-IL2 coding region and 3' moloney murine leukemia virus
XX CC (MoMuLV) LTR.
XX XX
XX SQ Sequence 4661 BP; 1257 A; 1096 C; 1161 G; 1147 T; 0 other;
XX XX
XX QY Query Match 1.1%; Score 31; DB 24; Length 4661;
XX QY Best Local Similarity 100.0%; Pred. No. 0.00012;
XX QY Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX XX
XX DB 4041 AATCTTTATTTTATCGATGTTAACAGCTT 4011
XX |
XX | 1 AATCTTTATTTTATCGATGTTAACAGCTT 31
XX |
XX XX
XX RESULT 49
XX AAD28312/C
XX ID AAD28312 standard; DNA; 4661 BP.
XX XX
XX AC AAD28312;
XX XX
XX DT 22-APR-2002 (first entry)
XX XX
XX DE Alpha-lactalbumin (LA) cc49IL2 vector.
XX XX
XX KM Regulatory element; vector; erythropoietin; growth hormone; insulin;
XX KM immunoglobulin; bone morphogenetic protein; interferon; interleukin;
XX KM superoxide dismutase; T-cell receptor; surface membrane protein;
XX KM viral antigen; transport protein; addressin; regulatory protein;

```


KM	bovine; alpha-lactalbumin; promoter; human; MoMuLV; chimeric;
KW	moloney murine leukemia virus; ds.
XX	
OS	Chimeric - Bos sp.
OS	Chimeric - Homo sapiens.
OS	Chimeric - Moloney murine leukemia virus.
OS	Chimeric - Unidentified.
XX	
FT	Key
FT	Location/Qualifiers
FT	misc_feature
FT	1..2055
FT	/*tag= a
FT	/note= "Bovine/human alpha-lactalbumin 5' flanking
FT	region"
FT	CDS
FT	2098..4011
FT	/*tag= b
FT	/product= "cc49-II2 coding region"
FT	4068..4661
FT	/*tag= c
FT	/note= "MoMuLV 3' LTR"
XX	
PN	WO200202783-A2.
PD	
PD	10-JAN-2002.
XX	
PF	29-JUN-2001; 2001MO-US20714.
PR	
PR	03-JUL-2000; 2000US-215851P.
PA	
PA	(GALA-) GALA DESIGN INC.
XX	
PI	Bleek GT;
XX	
DR	WPI; 2002-154749/20.
PT	
PT	Novel regulatory elements including nucleic acid encoding hybrid
PT	alpha-lactalbumin promoter or mutant RNA export element, for expressing
PT	one or more proteins e.g. antibodies, pharmaceutical proteins in host
PS	cells -
XX	
PS	Example 1; Fig 13; 151dp; English.
CC	The invention relates to novel regulatory elements and vectors for the
CC	expression of one or more proteins in a host cell. The invention further
CC	provides methods of indirectly detecting the expression of a protein of
CC	interest, comprising providing the host cell transfected with a vector
CC	encoding a polyelectronic sequence comprising a signal protein and a
CC	desired protein operably linked by an internal ribosome entry site
CC	(IRES), and culturing the host cell under suitable conditions so that
CC	the signal protein and the desired protein is produced, where the
CC	presence of signal protein indicates the presence of desired protein.
CC	Regulatory elements and vectors of the invention are useful for the
CC	expression of proteins of interest in a host cell. They are useful for
CC	producing an immunoglobulin (Ig), preferably secretory Ig. They are
CC	useful in the expression of one or more proteins such as erythropoietin,
CC	growth hormone, insulin, immunoglobulins, protein C, cytokines and their
CC	receptors, hormones, Von Willebrand factor, lung surfactant, serum
CC	albumins, DNase, vascular endothelial growth factor, receptors for
CC	hormones or growth factors, rheumatoid factors, nerve growth factors,
CC	CD proteins, osteoinductive factors, immunotoxins, bone morphogenetic
CC	protein, interferons, colony stimulating factors, interleukins,
CC	superoxide dismutase, T-cell receptors, surface membrane proteins,
CC	viral antigens, transport proteins, addressins, regulatory proteins,
CC	antibodies, chimeric proteins and their fragments. The vectors are
CC	particularly useful for expressing G protein coupled receptors and other
CC	transmembrane proteins. The retroviral vectors are useful for expressing
CC	proteins in mammalian tissue culture host cells, including rat fibroblast
CC	cells, bovine kidney cells and human kidney cells. The present sequence
CC	is alpha lactalbumin (IA) cc49II2 vector used in the invention. The
CC	vector comprises the following elements: bovine/human alpha-lactalbumin
CC	hybrid promoter; cc49-II2 coding region and LTR sequence from
CC	3' moloney murine leukemia virus (MoMuLV).
XQ	Sequence 4661 BP; 1257 A; 1096 C; 1161 G; 1147 T; 0 other;

Query Match	Similarity	1.1%;	Score 31;	DB 24;	Length 4661;
Best Local	Similarity	100.0%;	Pred. No. 0	00012;	
Matches	31;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Qy	1 AATCTTTATTTTATCGATGTTAAAGCTT	31			
Db	4041 AATCTTTATTTTATCGATGTTAAAGCTT	4011			
RESULT 50					
ABZ25421	standard; DNA; 29 BP.				
ABZ25421;					
ABZ25421;					
26-MAR-2003	(first entry)				
Human semaphorin-3A	PCR primer #5.				
Human; PCR; primer; cytosolic; tumour; semaphorin-3A;					
prostatic cancer; 5S.					
Homo sapiens.					
MO200286157-A2.					
31-OCT-2002.					
11-JAN-2002; 2002MO-FR00118.					
19-APR-2001; 2001US-284508P.					
(UROG-) UROGENE SA.					
Lac11 A, Quesnot O, Algarre-Genin M;					
WPI; 2003-093153/06.					
In vitro evaluation of invasiveness of tumor cells, useful for					
diagnosis and monitoring of prostatic cancer, comprises measuring					
expression of semaphorin-3A					
Disclosure; Page 17; 31pp; French.					
The present invention relates to a method for in vitro evaluation of the					
aggressiveness of tumor cells in a test sample. The method comprises					
measuring semaphorin-3A gene expression in both healthy and tumour					
epithelial cells. Cells are assessed as strongly invasive if the					
semaphorin-3A gene is underexpressed in them. The method is useful for					
diagnosing and evaluating the aggressiveness of prostatic cancer and to					
assess the efficacy of treatments. The present sequence is a PCR primer					
for semaphorin-3A, used in the method of the invention.					
Sequence 29 BP; 9 A; 6 C; 5 G; 9 T; 0 other;					
Query Match	1.1%;	Score 29;	DB 25;	Length 29;	
Best Local	Similarity	100.0%;	Pred. No. 0.0013;		
Matches	29;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
Qy	2155 GGAACATGGGTTTCATACAACTCTTCTTA	2183			
Db	1 GGAACATGGGTTTCATACAACTCTTCTTA	29			
RESULT 51					
AAF83093/C	standard; DNA; 9729 BP.				
AAF83093;					
29-JUN-2001	(first entry)				

```

DE Nucleotide sequence of MLV self-inactivating (SIN) vector pTRAP2.
DM
DM Retrovirus; recombinase recognition sequence; RRS; LTR; recombinase;
KM long terminal repeat; pharmaceutical; cytostatic; antiinflammatory;
KM antirheumatic; antialthritic; antialasthmatic; osteopathic; cardiac; MLV;
KM vasodilator; neuroprotective; nootropic; cerebroprotective; antipsoriatic;
KM antitartaric; sclerotic; vulnery; anti-HIV; anticancer; thrombolytic;
KM dermatological; gene therapy; ss.
KM
OS Synthetic.
OS Murine leukemia virus.
XX
XX WO200125466-A1.
XX
XX 12-APR-2001.
XX
XX 05-OCT-2000; 2000MO-GB03837.
XX
XX 05-OCT-1999; 99GB-0023558.
XX
XX (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
XX Slingsby J, Kingeman SM, Rohl J, Slade A;
XX WPI; 2001-281732/29.
XX
XX Modifying producer cells for making retrovirus by transfecting with a
XX construct comprising a 5'- recombinase recognition sequence, long
XX terminal repeat and 3'- recombinase recognition sequence, in presence
XX of recombinase -
XX
XX Example 6; Page 123-126; 133pp; English.
XX
XX The invention relates to a method of modifying producer cells for making
XX retrovirus by transfecting with a construct comprising a 5'- recombinase
XX recognition sequence (RRS), long terminal repeat (LTR) and 3'- RRS, in
XX presence of recombinase. The regulated retroviral vector produced is
XX useful in the manufacture of a pharmaceutical composition to deliver a
XX NOI to a target site, and in the manufacture of a medicament for
XX diagnostic, therapeutic and/or medical applications. The recombinase
XX assisted method is useful for introducing regulated 3'-LTR into a derived
XX producer cell line to produce a high titer regulated retroviral vector.
XX The vector is useful in gene therapy for treating diseases like cancers,
XX inflammatory diseases, immunological disorders such as graft vs host
XX disease, autoimmune diseases such as rheumatoid arthritis, allergic
XX diseases such as asthma, osteoporosis, cardiovascular diseases such as
XX congestive heart failure and ischemic heart disease, neurodegenerative
XX disorders such as multiple sclerosis, Alzheimer's disease, stroke and
XX cerebral ischemia, atherosclerosis, thrombotic disorders, dermatological
XX disorders such as atopic dermatitis, contact dermatitis and psoriasis,
XX wound healing, retinosis, infectious disorders such as HIV infections,
XX ulcers, digestive disorders such as anorexia, bulimia and cachexia, and
XX other diseases. The present sequence represents the nucleotide sequence
XX of a Murine leukemia virus (MLV) self-inactivating (SIN) vector pTRAP2.
XX
XX Sequence 9729 BP; 2270 A; 2657 C; 2540 G; 2262 T; 0 other;
XX
XX Query Match 1.0%; Score 27; DB 22; Length 9729;
XX Best Local Similarity 100.0%; Pred. No. 0.011;
XX Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
XX
XX 1 AATCTTTATTTATCGATGTTAAACA 27
XX ||||||||||||||||||||
XX Db 5278 AATCTTTATTTATCGATGTTAAACA 5252
XX
XX RESULT 52
XX AAF83092/C
XX ID AAF83092 standard; DNA; 12473 BP.
XX
XX .AAF83092;
XX
XX 29-JUN-2001 (first entry)
XX

```

XX		Nucleotide sequence of MLV construct CZCS.
DE		
XX		
KW	Retrovirus; recombinase recognition sequence; RRS; LTR; recombinase;	
KM	long terminal repeat; pharmaceutical; cytostatic; antiinflammatory;	
KX	antihypertensive; antidiabetic; antiaspartic; osteopathic; cardiac;	
KW	vaccinogenic; neuroprotective; nootropic; cerebroprotective; antisoriatic;	
KM	antiarteriosclerotic; vlnarary; anti-HIV; anticancer; thrombolytic;	
KW	dermatological; gene therapy; ss.	
XX		
OS	Synthetic.	
OS	Murine leukemia virus.	
XN		
PN	WO200125466-AI.	
PD		
XX	12-APR-2001.	
PF		
PP	05-OCT-2000; 2000WO-GB03837.	
PR		
XX	05-OCT-1999; 99GB-0023558.	
PA	(OXFO-) OXFORD BIOMEDICA UK LTD.	
DR		
PI	Slingsby J, Kingsman SM, Rohll J, Slade A;	
XX	WPI; 2001-281732/29.	
PT		
PT	Modifying producer cells for making retrovirus by transfecting with a	
PT	construct comprising a 5'-recombinase recognition sequence, long	
PT	terminal repeat and 3'-recombinase recognition sequence, in presence	
PT	of recombinase -	
PS		
XX	Examples; Page 120-123; 133pp; English.	
CC	The invention relates to a method of modifying producer cells for making	
CC	retrovirus by transfecting with a construct comprising a 5'-recombinase	
CC	recognition sequence (RRS), long terminal repeat (LTR) and 3'-RRS, in	
CC	presence of recombinase. The regulated retroviral vector produced is	
CC	useful in the manufacture of a pharmaceutical composition to deliver a	
CC	NOI to a target site, and in the manufacture of a medicament for	
CC	diagnostic, therapeutic and/or medical applications. The recombinase	
CC	assisted method is useful for introducing regulated 3'-LTR into a derived	
CC	producer cell line to produce a high titer regulated retroviral vector.	
CC	The vector is useful in gene therapy for treating diseases like cancers,	
CC	inflammatory diseases, immunological disorders such as graft vs host	
CC	disease, autoimmune diseases such as rheumatoid arthritis, allergic	
CC	diseases such as asthma, osteoporosis, cardiovascular diseases such as	
CC	congestive heart failure and ischemic heart disease, neurodegenerative	
CC	disorders such as multiple sclerosis, Alzheimer's disease, stroke and	
CC	cerebral ischemia, atherosclerosis, thrombotic disorders, dermatological	
CC	disorders such as atopic dermatitis, contact dermatitis and psoriasis,	
CC	wound healing, restenosis, infectious disorders such as HIV infections,	
CC	ulcers, digestive disorders such as anorexia, bulimia and cachexia, and	
CC	other diseases. The present sequence represents the nucleotide sequence	
CC	of a Murine leukemia virus (MLV) construct CZCS.	
XX		
XX		
BQ	Sequence 12473 BP; 2841 A; 3434 C; 3357 G; 2841 T; 0 other;	
Query Match	1.0%; Score 27; DB 22; Length 12473;	
Best Local Similarity	100.0%; Pred. No. 0.011;	
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0		
CY	1 AAATGTTTATTTCAGTGTTAACAA 27 	
DB	5278 AATCCTTAATTTCAGTGTTAACAA 5252	
RESULT 53		
ID	AAS2430/C	
XX	AAS2430 standard; CDNA, 1313 BP.	
AC	AAS2430;	

DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 975.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorder; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200157190-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US04098.
XX
PR 03-FEB-2000; 2000US-0496914.
PR 27-APR-2000; 2000US-0560875.
PR 20-JUN-2000; 2000US-0598075.
PR 19-JUL-2000; 2000US-0620325.
PR 01-SEP-2000; 2000US-0654936.
PR 15-SEP-2000; 2000US-0663561.
PR 20-OCT-2000; 2000US-0693325.
PR 30-NOV-2000; 2000US-0728422.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Tang YF, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Gao Y, Ma Y;
PI Zhao Q, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;
XX
XX WPI; 2001-476283/51.
DR P-PSDB; AAM79297.
XX
PT Nucleic acids encoding polypeptides with cytokine-like activities,
PT useful in diagnosis and gene therapy -
XX
XX Claim 1; Page 3186-3187; 6221pp; English.
XX
XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM79323-AAW80302) that exhibit actively elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
CC (AAM80020) are omitted as the relevant pages from the sequence listing
CC were missing at the time of publication.
XX
SQ Sequence 1313 BP; 341 A; 304 C; 241 G; 427 T; 0 other;
XX
XX Query Match 1.0%; Score 26; DB 22; Length 1313;
XX Best Local Similarity 100.0%; Pred. No. 0.037;
XX Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 55 GTGTCGACCAACGCGTCCGGAGTAG 80
DB 1305 GTGTCGACCAACGCGTCCGGAGTAG 1280
XX
XX RESULT 54
XX ABZ25416/C
XX ID ABZ25416 standard; DNA; 25 BP.
XX AC ABZ25416;
XX XX
XX 26-MAR-2003 (first entry)

XX
XX Human semaphorin-3A PCR primer #2.
DE
XX
XX Human; PCR; primer; cytosolic; tumour; semaphorin-3A;
KW prostatic cancer; ss.
XX
XX
OS Homo sapiens.
XX
PN WO200286157-A2.
XX
PD 31-OCT-2002.
XX
XX 11-JAN-2002; 2002WO-FR00118.
XX
XX 19-APR-2001; 2001US-284508P.
XX
XX (UROG-) UROGENE SA.
XX
XX Latil A, Cussenot O, Algarre-Genin M;
PI WPI; 2003-093153/08.
XX
XX
XX In vitro evaluation of invasiveness of tumor cells, useful for
PT diagnosis and monitoring of prostatic cancer, comprises measuring
PT expression of semaphorin-3A -
XX
XX
XX Disclosure; Page 16; 31pp; French.
XX
XX The present invention relates to a method for in vitro evaluation of the
CC aggressiveness of tumor cells in a test sample. The method comprises
CC measuring semaphorin-3A gene expression in both healthy and tumour
CC epithelial cells. Cells are assessed as strongly invasive if the
CC semaphorin-3A gene is underexpressed in them. The method is useful for
CC diagnosing and evaluating the aggressiveness of prostatic cancer and to
CC assess the efficacy of treatments. The present sequence is a PCR primer
CC for semaphorin-3A, used in the method of the invention.
XX
SQ Sequence 25 BP; 7 A; 7 C; 4 G; 7 T; 0 other;
XX
XX Query Match 0.9%; Score 25; DB 25; Length 25;
XX Best Local Similarity 100.0%; Pred. No. 0.13;
XX Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1556 ACAGATGTTGGACCGTCTTAAG 1580
DB 25 ACAGATGTTGGACCGTCTTAAG 1
XX
XX RESULT 55
XX AAH47052
XX ID AAH47052 standard; DNA; 1399 BP.
XX AC AAH47052;
XX
XX 29-OCT-2001 (first entry)
XX
XX Murine maspin cDNA sequence.
DE
XX
XX Drug resistance; resistance gene; semaphorin D; B94; mel-14 antigen;
KW 24p3; proliferin; maspin; cancer; cytosolic; gene therapy; ss.
XX
XX Mus sp.
XX
XX WO200155455-A2.
XX
XX 02-AUG-2001.
XX
XX 31-JAN-2001; 2001WO-US03161.
XX
XX 31-JAN-2000; 2000US-0179191.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX (JINS/) JIN S.

XX Jtn S;
PI
XX
XX
DR WPI; 2001-488799/53.
XX
PT Determining if a compound modulates the drug resistance of a cell,
PT comprises determining the expression or activity level of a resistance
PT sequence in a cell in the presence of the test compound -
XX
XX Example 1; Fig 4; 79pp; English.
XX
XX The invention relates to a method of determining whether a test compound
CC modulates the drug resistance of a cell that comprises determining the
CC expression or activity level of resistance genes (e.g. semaphorin D, B94,
CC mel-14 antigen, 24p3, proliferin or maspin) in a cell in the presence of
CC the test compound, and comparing its expression or activity level in a
CC cell without the test compound. The drug resistant sequences are useful
CC in identifying drug resistant cells, in screening methods directed to the
CC identification of compounds that can modulate the drug resistance of a
CC cell type or multiple cell types. An isolated resistance protein can be
CC used as an immunogen to generate antibodies that bind the resistance
CC protein. Resistance nucleic acids may be inserted into vectors and used
CC as gene therapy vectors. An anti-resistance protein antibody may be used
CC to isolate a resistance protein, or facilitate the purification of
CC natural resistance protein from cells and of recombinantly produced
CC resistance protein expressed in host cells. The methods are useful for
CC treating a subject having a disorder, such as a drug-resistance cancer,
CC characterized by aberrant resistance sequence expression or activity by
CC administering to the subject a resistance sequence modulator. The present sequence
CC represents a murine maspin cDNA sequence, whose expression was decreased
CC in drug resistant EMT6 tumours.
XX
XX
SQ Sequence 1399 BP; 436 A; 313 C; 278 G; 372 T; 0 other;
Query Match 0.9%; Score 25; DB 22; Length 1399;
Best Local Similarity 100.0%; Pred. No. 0.12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 22 TAACAAGCTTAGTATGATGCCAC 46
DB 29 TAACAAGCTTAGTATGATGCCAC 53
RESULT 56
AB225415
ID AB225415 standard; DNA; 24 BP.
XX
XX AB225415;
AC
XX
XX 26-MAR-2003 (first entry)
DT
XX
XX Human semaphorin-3A PCR primer #1.
DE
XX
XX Human; PCR; primer; cytosolic; tumour; semaphorin-3A;
KM prostatic cancer; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200286157-A2.
PN
XX
XX 31-OCT-2002.
PD
XX
XX 11-JAN-2002; 2002WO-FR00118.
PF
XX
XX 19-APR-2001; 2001US-284508P.
PR
XX
XX (UROG-) UROGENE SA.
PA
XX
XX Latil A, Cussenot O, Algarde-Genin M;
PI
XX
XX WPI; 2003-093153/08.
DR
XX
XX In vitro evaluation of invasiveness of tumor cells, useful for

PT diagnosis and monitoring of prostatic cancer, comprises measuring
PT expression of semaphorin-3A -
XX
XX
XX Disclosure; Page 16; 31pp; French.
PS
XX
XX The present invention relates to a method for in vitro evaluation of the
CC aggressiveness of tumour cells in a test sample. The method comprises
CC measuring semaphorin-3A gene expression in both healthy and tumour
CC epithelial cells. Cells are assessed as strongly invasive if the
CC semaphorin-3A gene is underexpressed in them. The method is useful for
CC diagnosing and evaluating the aggressiveness of prostatic cancer and to
CC assess the efficacy of treatments. The present sequence is a PCR primer
CC for semaphorin-3A, used in the method of the invention.
XX
XX
SQ Sequence 24 BP; 8 A; 7 C; 4 G; 5 T; 0 other;
Query Match 0.9%; Score 24; DB 25; Length 24;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1442 CCTATGAACAATCGCCCAATAGTG 1465
DB 1 CCTATGAACAATCGCCCAATAGTG 24
RESULT 57
AAC66819/C
ID AAC66819 standard; DNA; 8417 BP.
XX
XX AAC66819;
AC
XX
XX 27-FEB-2001 (first entry)
DT
XX
XX Reporter gene construct vector pRTRapsin-TK/BSN3.5'.
DE
XX
XX Reporter gene construct; gene trap; gene function; biological pathway;
KM isotype switching; drug identification; ds.
KW
XX
XX Synthetic.
OS
XX
XX WO200061809-A2.
PN
XX
XX 19-OCT-2000.
PD
XX
XX 07-APR-2000; 2000WO-US09267.
PF
XX
XX 09-APR-1999; 99US-0128631.
PR
XX
XX (ICON-) ICONIX PHARM INC.
PA
XX
XX Xu H, Jarnigan K, Zhou H, Greene A, Thode S;
PI
XX
XX WPI; 2000-679501/66.
DR
XX
XX Reporter constructs for assaying biological response modifiers,
PT comprises a splice acceptor sequence, a first element comprising three
PT subelements, a reporter sequence, and a positive and negative
PT selectable marker -
XX
XX
XX Disclosure; Page 35-37; 57pp; English.
PS
XX
XX The present invention relates to novel reporter gene constructs,
CC designated gene traps, comprising a splice acceptor sequence, a reporter
CC sequence and positive and negative selectable markers. They can be used
CC in the expression of heterologous genes, to determine gene function, in
CC biological pathway analysis, to understand the mechanisms of isotype
CC switching and to identify drugs, particularly allergy drugs which
CC interfere with IgE production. The present sequence is one of the
CC reporter gene constructs of the invention.
XX
XX
SQ Sequence 8417 BP; 1892 A; 2354 C; 2199 G; 1972 T; 0 other;
Query Match 0.9%; Score 24; DB 21; Length 8417;

Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AATCTTTATTTATCGATGTTAA 24
Db 5703 AATCTTTATTTATCGATGTTAA 5680

RESULT 58
AAC66820/c
ID AAC66820 standard; DNA; 8684 BP.
XX
AC AAC66820;
XX
DT 27-FEB-2001 (first entry)
XX
DE Reporter gene construct vector pRTrip-TK/BSN3.5'.
XX
KW Reporter gene construct; gene trap; gene function; biological pathway;
KW isotype switching; drug identification; ds.
XX
OS Synthetic.
XX
PM WO200061809-A2.
XX
PD 19-OCT-2000.
XX
PF 07-APR-2000; 2000WO-US09267.
XX
PR 09-APR-1999; 99US-0128631.
XX
PA (ICON-) ICONIX PHARM INC.
XX
PI Xu H, Jarnigan K, Zhou H, Greene A, Thode S;
XX
DR WPI; 2000-679501/66.
XX
PT Reporter constructs for assaying biological response modifiers,
PT comprises a splice acceptor sequence, a first element comprising three
PT subelements, a reporter sequence, and a positive and negative
PT selectable marker -
XX
PS Disclosure; Page 38-40; 57PD; English.
XX
CC The present invention relates to novel reporter gene constructs,
CC designated gene traps, comprising a splice acceptor sequence, a reporter
CC sequence and positive and negative selectable markers. They can be used
CC in the expression of heterologous genes, to determine gene function, in
CC biological pathway analysis, to understand the mechanisms of isotype
CC switching and to identify drugs, particularly allergy drugs which
CC interfere with IgE production. The present sequence is one of the
CC reporter gene constructs of the invention.
XX
SQ Sequence 8684 BP; 1975 A; 2415 C; 2271 G; 2023 T; 0 other;
Query Match 0.9%; Score 24; DB 21; Length 8684;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 AATCTTTATTTATCGATGTTAA 24
Db 5703 AATCTTTATTTATCGATGTTAA 5680

RESULT 59
AB225419
ID AB225419 standard; DNA; 29 BP.
XX
AC AB225419;
XX
DT 26-MAR-2003 (first entry)
XX
DE Human semaphorin-3A PCR primer #3.

XX
KW Human; PCR; primer; cytoskeletal; tumour; semaphorin-3A;
KW prostatic cancer; ss.
XX
OS Homo sapiens.
XX
PM WO200286157-A2.
XX
PD 31-OCT-2002.
XX
PF 11-JAN-2002; 2002WO-FR00118.
XX
PR 19-APR-2001; 2001US-284508P.
XX
PA (UROG-) UROGENE SA.
XX
PI Latil A, Cussenot O, Algarre-Genin M;
XX
DR WPI; 2003-093153/08.
XX

PT In vitro evaluation of invasiveness of tumor cells, useful for
PT diagnosis and monitoring of prostatic cancer, comprises measuring
PT expression of semaphorin-3A -
XX
PS Disclosure; Page 17; 31PD; French.
XX

CC The present invention relates to a method for in vitro evaluation of the
CC aggressiveness of tumor cells in a test sample. The method comprises
CC measuring semaphorin-3A gene expression in both healthy and tumour
CC epithelial cells. Cells are assessed as strongly invasive if the
CC semaphorin-3A gene is underexpressed in them. The method is useful for
CC diagnosing and evaluating the aggressiveness of prostatic cancer and to
CC assess the efficacy of treatments. The present sequence is a PCR primer
CC for semaphorin-3A, used in the method of the invention.
XX
SQ Sequence 29 BP; 6 A; 6 C; 9 G; 8 T; 0 other;
Query Match 0.8%; Score 23; DB 25; Length 29;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 192 TCTGACGATGGCTGTTAACT 214
Db 7 TCTGACGATGGCTGTTAACT 29

RESULT 60
ABK87383/c
ID ABK87383 standard; CDNA; 175 BP.
XX
AC ABK87383;
XX
DT 24-SEP-2002 (first entry)
XX
DE CDNA fragment for pertubagen R2.
XX

KW Retinoid acid pathway; RA; ss; retinoid; lung cancer;
KW Kaposi's sarcoma; breast cancer; pancreatic cancer; neuroblastoma;
KW renal cancer; ovarian cancer; dermatitis; hyperkeratosis; eczema;
KW Darier's disease; Reiter's disease; psoriasis; gene therapy; R2;
KW acute promyelocytic leukaemia; APL; Pertubagen; R3; F802; F820;
KW yeast two-hybrid assay; PAT1; kinesin light chain-related protein;
KW human; EGFP; green fluorescent protein.
XX
OS Homo sapiens.
OS Aequorea victoria.
OS Synthetic.
XX
PM WO200240719-A2.
XX
PD 23-MAY-2002.
XX
PF 17-NOV-2001; 2001WO-US44039.

XX 17-NOV-2000; 2000US-249468P.
 XX (DELTA-) DELTAGEN PROTEOMICS INC.
 PA Kamb CA, Richards BT, Karpilow J;
 PI WPI; 2002-519386/55.
 XX Polypeptide with retinoic acid pathway activity, especially of
 PT perturbagens R3, F802 and F820 for identifying a cellular target which
 PT interacts with the polypeptide and for therapeutic purposes
 XX Example 4; Fig 15a; 131pp; English.
 XX The invention relates to an isolated polypeptide (I) with retinoic acid
 CC (RA) pathway activity, comprising a polypeptide sequence of Perturbagen
 CC (phenotypic probe) R3, Perturbagen F802, Perturbagen F820, their
 CC biologically active modifications, or biologically active fragments.
 CC Also included are the polynucleotides encoding the perturbagens, a
 CC vector comprising the polynucleotide, preparing an RA pathway related
 CC polypeptide, a composition comprising the polypeptide, an antibody to the
 CC polypeptide, screening (M) putative RA-related therapeutics, by exposing
 CC a polypeptide/target interaction pair identified by a method utilizing
 CC the RA pathway polypeptide, to a number of agents, and recovering a
 CC subpopulation of disrupting agents which competitively displace the
 CC polypeptide from the target, where the disrupting agents are putative RA-
 CC related therapeutics. Also include are an isolated RA pathway polypeptide
 CC comprising PARI (a kinesin light chain-related protein) polypeptide
 CC and its encoding polynucleotide, a gene therapy vector comprising the RA
 CC pathway protein polypeptide or encoding or PARI and a host cell
 CC comprising the gene therapy vector. The RA pathway polypeptide is useful
 CC for identifying a cellular target that interacts with RA pathway-related
 CC polypeptide, by exposing the polypeptide in vitro to putative target
 CC molecules and identifying a polypeptide/target interaction pair, by
 CC detecting reporter expression, where the reporter expression is
 CC operatively linked to the formation of the interaction pair. (M) is a
 CC yeast two-hybrid assay. The polypeptide is also useful for treating
 CC an RA pathway-related condition e.g. lung cancer, Kaposi's sarcoma,
 CC breast cancer, pancreatic cancer, neuroblastoma, renal cancer,
 CC ovarian cancer, dermatitis, hyperkeratosis, eczema, Darier's disease,
 CC Reiter's disease, psoriasis, acute promyelocytic leukaemia (APL).
 CC The present sequence is a cDNA fragment representing perturbagen R2
 CC and containing sequences derived from the vector (EGFP, inactive green
 CC fluorescent protein sequence).
 XX
 XX Sequence 175 BP; 40 A; 34 C; 60 G; 41 T; 0 other;
 SQ
 Query Match 0.8%; Score 23; DB 24; Length 175;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AATCTTTATTTATCGATGTTA 23
 |||
 Db 169 AATCTTTATTTATCGATGTTA 147
 |||
 RESULT 61
 ABK87385/c
 ID ABK87385 standard; cDNA, 240 BP.
 XX
 AC ABK87385;
 XX
 DT 24-SEP-2002 (first entry)
 XX
 DE cDNA fragment for perturbagen R44.
 XX
 XX Retinoic acid pathway; RA; ss; retinoid; lung cancer;
 KM Kaposi's sarcoma; breast cancer; pancreatic cancer; neuroblastoma;
 KM renal cancer; ovarian cancer; dermatitis; hyperkeratosis; eczema;
 KM Darier's disease; Reiter's disease; psoriasis; gene therapy; R44;
 KM acute promyelocytic leukaemia; APL; Perturbagen; R3; F802; F820;
 KM yeast two-hybrid assay; PARI; kinesin light chain-related protein;

KM human; EGFP; green fluorescent protein.
 XX
 XX Homo sapiens.
 OS Aegreora victorica.
 OS Synthetic.
 XX
 PN WO200240719-A2.
 XX
 PD 23-MAY-2002.
 XX
 PF 17-NOV-2001; 2001WO-US44039.
 XX
 PR 17-NOV-2000; 2000US-249468P.
 XX
 PA (DELTA-) DELTAGEN PROTEOMICS INC.
 PI Kamb CA, Richards BT, Karpilow J;
 XX WPI; 2002-519386/55.
 XX Polypeptide with retinoic acid pathway activity, especially of
 PT perturbagens R3, F802 and F820 for identifying a cellular target which
 PT interacts with the polypeptide and for therapeutic purposes
 XX Example 4; Fig 15b; 131pp; English.
 XX The invention relates to an isolated polypeptide (I) with retinoic acid
 CC (RA) pathway activity, comprising a polypeptide sequence of Perturbagen
 CC (phenotypic probe) R3, Perturbagen F802, Perturbagen F820, their
 CC biologically active modifications, or biologically active fragments.
 CC Also included are the polynucleotides encoding the perturbagens, a
 CC vector comprising the polynucleotide, preparing an RA pathway related
 CC polypeptide, a composition comprising the polypeptide, an antibody to the
 CC polypeptide, screening (M) putative RA-related therapeutics, by exposing
 CC a polypeptide/target interaction pair identified by a method utilizing
 CC the RA pathway polypeptide, to a number of agents, and recovering a
 CC subpopulation of disrupting agents which competitively displace the
 CC polypeptide from the target, where the disrupting agents are putative RA-
 CC related therapeutics. Also include are an isolated RA pathway polypeptide
 CC comprising PARI (a kinesin light chain-related protein) polypeptide
 CC and its encoding polynucleotide, a gene therapy vector comprising the RA
 CC pathway protein polypeptide or encoding or PARI and a host cell
 CC comprising the gene therapy vector. The RA pathway polypeptide is useful
 CC for identifying a cellular target that interacts with RA pathway-related
 CC polypeptide, by exposing the polypeptide in vitro to putative target
 CC molecules and identifying a polypeptide/target interaction pair, by
 CC detecting reporter expression, where the reporter expression is
 CC operatively linked to the formation of the interaction pair. (M) is a
 CC yeast two-hybrid assay. The polypeptide is also useful for treating
 CC an RA pathway-related condition e.g. lung cancer, Kaposi's sarcoma,
 CC breast cancer, pancreatic cancer, neuroblastoma, renal cancer,
 CC ovarian cancer, dermatitis, hyperkeratosis, eczema, Darier's disease,
 CC Reiter's disease, psoriasis, acute promyelocytic leukaemia (APL).
 CC The present sequence is a cDNA fragment representing perturbagen R44
 CC and containing sequences derived from the vector (EGFP, inactive green
 CC fluorescent protein sequence).
 XX
 XX Sequence 240 BP; 65 A; 32 C; 87 G; 56 T; 0 other;
 SQ
 Query Match 0.8%; Score 23; DB 24; Length 240;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 AATCTTTATTTATCGATGTTA 23
 |||
 Db 234 AATCTTTATTTATCGATGTTA 212
 |||
 RESULT 62
 ABK87379/c
 ID ABK87379 standard; cDNA, 287 BP.
 XX
 AC ABK87379;

XX 24-SEP-2002 (first entry)
 DT CDNA encoding partial perturbation F802.
 XX
 DE Retinoic acid pathway; RA; ss; retinoid; lung cancer;
 XX Kaposi's sarcoma; breast cancer; pancreatic cancer; neuroblastoma;
 KW renal cancer; ovarian cancer; dermatitis; hyperkeratosis; eczema;
 KW Darier's disease; Reiter's disease; psoriasis; gene therapy; F797;
 KW acute promyelocytic leukemia; APL; Perturbation; R3; F802; F820;
 KW yeast two-hybrid assay; PARI; kinesin light chain-related protein;
 KW human; EGFP; green fluorescent protein.
 XX Homo sapiens.
 OS Aegoreia victoria.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT CDS /tag= a
 FT 44..238 /product= "Partial perturbation F802"
 FT /partial
 FT /note= "No start codon shown"
 FT
 FT
 FT
 XX WO200240719-A2.
 XX 23-MAY-2002.
 XX 17-NOV-2001; 2001WO-US44039.
 XX 17-NOV-2000; 2000US-249468P.
 XX (DELT-) DELTAGEN PROTEOMICS INC.
 XX
 XX Kamb CA, Richards BT, Karpilow J;
 FI WPI; 2002-519386/55.
 DR P-PSDB; AAU99181.
 XX
 PT Polypeptide with retinoic acid pathway activity, especially of
 PT perturbagens R3, F802 and F820 for identifying a cellular target which
 PT interacts with the polypeptide and for therapeutic purposes -
 XX
 PS Claim 67; Fig 13a; 131pp; English.
 XX
 CC The invention relates to an isolated polypeptide (I) with retinoic acid
 CC (RA) pathway activity, comprising a polypeptide sequence of Perturbation
 CC (phenotypic probe) R3, Perturbation F802, Perturbation F820, their
 CC biologically active modifications, or biologically active fragments.
 CC Also included are the polynucleotides encoding the perturbagens, a
 CC vector comprising the polynucleotide, preparing an RA pathway related
 CC polypeptide, a composition comprising the polypeptide, an antibody to the
 CC polypeptide, screening (M) putative RA-related therapeutics, by exposing
 CC a polypeptide/target interaction pair identified by a method utilising
 CC the RA pathway polypeptide, to a number of agents, and recovering a
 CC subpopulation of disrupting agents which competitively displace the
 CC polypeptide from the target, where the disrupting agents are putative RA-
 CC related therapeutics. Also include are an isolated RA pathway polypeptide
 CC comprising PARI (a kinesin light chain-related protein) polypeptide
 CC and its encoding polynucleotide, a gene therapy vector comprising the RA
 CC pathway protein polypeptide or encoding or PARI and a host cell
 CC comprising the gene therapy vector. The RA pathway polypeptide is useful
 CC for identifying a cellular target that interacts with RA pathway-related
 CC polypeptide, by exposing the polypeptide in vitro to putative target
 CC molecules and identifying a polypeptide/target interaction pair, by
 CC detecting reporter expression, where the reporter expression is
 CC operatively linked to the formation of the interaction pair. (M) is a
 CC yeast two-hybrid assay. The polypeptide is also useful for treating
 CC an RA pathway-related condition e.g. lung cancer, Kaposi's sarcoma,
 CC breast cancer, pancreatic cancer, neuroblastoma, renal cancer,
 CC ovarian cancer, dermatitis, hyperkeratosis, eczema, Darier's disease,
 CC Reiter's disease, psoriasis, acute promyelocytic leukemia (APL).
 CC The present sequence is a CDNA encoding partial perturbation F802

CC and containing sequences derived from the vector (EGFP, inactive green
 CC fluorescent protein sequence).
 CC
 XX
 SQ Sequence 287 BP; 93 A; 61 C; 74 G; 59 T; 0 other;

Query Match 0.8%; Score 23; DB 24; Length 287;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Gy 1 AATCTTTATTTATCGATGTTA 23
 Db 280 AATCTTTATTTATCGATGTTA 258

RESULT 63

AA18251/c
 ID AA18251 standard; CDNA; 326 BP.

AA18251;

DT 07-DEC-2001 (first entry)

DE Human breast cancer expressed polynucleotide 10708.

XX Human; breast cancer; cell marker; cytostatic; ss.

XX Homo sapiens.

XX WO200151628-A2.

XX 19-JUL-2001.

XX 10-JAN-2001; 2001WO-US00798.

XX 14-JAN-2000; 2000US-0176077.

XX 14-MAR-2000; 2000US-0189167.

XX 24-MAR-2000; 2000US-0192099.

XX 29-MAR-2000; 2000US-0193480.

XX 15-MAY-2000; 2000US-0205230.

XX 09-JUN-2000; 2000US-0211315.

XX 25-JUL-2000; 2000US-0220534.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Lillie J, Xu Y, Wang Y, Steinmann K;

XX WPI; 2001-451856/48.

XX New peptide useful as a marker for the diagnosis of breast cancer -

XX Claim 1; Page 1911; 3695pp; English.

XX The invention relates to human breast cancer expressed polynucleotides

XX (AAU07544-AAU26789) and methods of assessing whether a patient is

XX afflicted with breast cancer by examining the correlation between the

XX expression of certain markers and the cancerous state of breast cells.

XX The polynucleotides and encoded polypeptides are potential markers for

XX detecting, diagnosing, monitoring, characterizing treating and

XX potentially preventing breast cancer. The polynucleotides and encoded

XX polypeptides are also useful for isolating compounds with cytostatic

XX activity.

XX Sequence 326 BP; 73 A; 51 C; 79 G; 123 T; 0 other;

Query Match 0.8%; Score 23; DB 22; Length 326;

Best Local Similarity 100.0%; Pred. No. 1.2;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Gy 54 GGTGTGACCCACCGGTCGGGA 76

Db 294 GGTGTGACCCACCGGTCGGGA 272

RESULT 64	ABK87384/c
ID	ABK87384 standard; cDNA; 335 BP.
XX	
XX	
XX	ABK87384;
DT	24-SEP-2002 (first entry)
XX	
DE	cDNA fragment for perturbation R3.
XX	
KW	Retinoic acid pathway; RA; ss; retinoid; lung cancer;
KW	Kapost's sarcoma; breast cancer; pancreatic cancer; neuroblastoma;
KW	renal cancer; ovarian cancer; dermatitis; hyperkeratosis; eczema;
KW	Darier's disease; Reiter's disease; psoriasis; gene therapy; R3;
KW	acute promyelocytic leukemia; APL; Perturbation; R3; F802; F820;
KW	yeast two-hybrid assay; PAT1; kinesin light chain-related protein;
KW	human; EGFP; green fluorescent protein.
XX	
OS	Homo sapiens.
OS	Aegoreta victoria.
XX	
XX	Synthetic.
XX	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	44..73
FT	/tag= a
FT	/partial
FT	/product= "Partial perturbation R3"
FT	/note= "No start codon shown"
PN	
MO	MO200240719-A2.
PD	
XX	23-MAY-2002.
XX	
XX	17-NOV-2001; 2001WO-US44039.
XX	
XX	17-NOV-2000; 2000US-249468P.
XX	
PA	(DELTA-) DELTAGEN PROTEOMICS INC.
PI	
PI	Kamb CA, Richards BT, Karpilow J;
XX	
XX	WPI; 2002-519386/55.
DR	P-PSDB; AAU99184.
DR	
XX	
PT	Polypeptide with retinoic acid pathway activity, especially of
PT	perturbagens R3, F802 and F820 for identifying a cellular target which
PT	interacts with the polypeptide and for therapeutic purposes -
PS	Claim 67; Fig 15a; 131pp; English.
CC	
CC	The invention relates to an isolated polypeptide (I) with retinoic acid
CC	(RA) pathway activity, comprising a polypeptide sequence of Perturbagen
CC	(phenotypic probe) R3, Perturbagen F802, Perturbagen F820, their
CC	biologically active modifications, or biologically active fragments.
CC	Also included are the polynucleotides encoding the perturbagens, a
CC	vector comprising the polynucleotide, preparing an RA pathway related
CC	polypeptide, a composition comprising the polypeptide, an antibody to the
CC	polypeptide, screening (M) putative RA-related therapeutics, by exposing
CC	a polypeptide/target interaction pair identified by a method utilising
CC	the RA pathway polypeptide, to a number of agents, and recovering a
CC	subpopulation of disrupting agents which competitively displace the
CC	polypeptide from the target, where the disrupting agents are putative RA-
CC	related therapeutics. Also include are an isolated RA pathway polypeptide
CC	comprising PAT1 (a kinesin light chain-related protein) polypeptide
CC	and its encoding polynucleotide, a gene therapy vector comprising the RA
CC	pathway protein polypeptide or encoding or PAT1 and a host cell
CC	comprising the gene therapy vector. The RA pathway polypeptide is useful
CC	for identifying a cellular target that interacts with RA pathway-related
CC	polypeptide, by exposing the polypeptide in vitro to putative target
CC	molecules and identifying a polypeptide/target interaction pair, by
CC	detecting reporter expression, where the reporter expression is
CC	operatively linked to the formation of the interaction pair. (M) is a
CC	yeast two-hybrid assay. The polypeptide is also useful for treating

CC	an RA pathway-related condition e.g. lung cancer, Karpov's sarcoma,
CC	broad cancer, pancreatic cancer, neuroblastoma, renal cancer,
CC	ovarian cancer, dermatitis, hyperkeratosis, eczema, Darier's disease,
CC	Richter's disease, psoriasis, acute promyelocytic leukaemia (APL).
CC	The present sequence is a cDNA fragment representing perturbation R3
CC	and containing sequences derived from the vector (EGFP, inactive green
XX	fluorescent protein sequence).
SO	Sequence 335 BP; 112 A; 43 C; 84 G; 96 T; 0 other;
Query Match	0.8%; Score 23; DB 24; Length 335;
Best Local Similarity	100.0%; Pred. No. 1.2;
Matches 23; Conservative	0; Mismatches 0; Indels 0; Gaps
Oy	1 AATCTTTATTTCGATGTA 23 329 AATCTTTATTTCGATGTA 307
Db	
RESULT 65	
ABK87380/c	
ID	ABK87380 standard; cDNA; 339 BP.
XX	
AC	ABK87380;
XX	
DT	24-SEP-2002 (first entry)
DE	cDNA encoding partial perturbation F820.
XX	
KW	Retinoid acid pathway; RA; ss; retinoid; lung cancer;
KW	Karpov's sarcoma; breast cancer; pancreatic cancer; neuroblastoma;
KW	renal cancer; ovarian cancer; dermatitis; hyperkeratosis; eczema;
KW	Darier's disease; Richter's disease; psoriasis; gene therapy; F797;
KW	acute promyelocytic leukaemia; APL; Perturbation; R3; F802; F820;
KW	yeast two-hybrid assay; PAPI; kinesin light chain-related protein;
XX	human; EGFP; green fluorescent protein.
OS	Homo sapiens.
OS	Aequorea victoria.
OS	Synthetic.
XX	
FH	Location/Qualifiers
FT	CDS
FT	44..136
FT	/tag= a
FT	/product= "Partial perturbation F820"
FT	/partial
FT	/note= "No start codon shown"
FT	/transl_except= (pos:56..61,aa:Ile)
XX	
FN	WO200240719-A2.
XX	
PD	23-MAY-2002.
XX	
PF	17-NOV-2001; 2001WO-US44039.
XX	
PR	17-NOV-2000; 2000US-249468P.
XX	
PA	(DELTA-) DELTAGEN PROTEOMICS INC.
PI	Kamb CA, Richards BT, Karpilow J;
XX	
DR	WPI; 2002-519386/55.
DR	P-PSDB; AAU99182.
XX	
PT	Polypeptide with retinoid acid pathway activity, especially of
PT	perturbagens R3, F802 and F820 for identifying a cellular target which
XX	interacts with the polypeptide and for therapeutic purposes -
XX	
PS	Claim 67; Fig 13b; 131pp; English.
XX	
CC	The invention relates to an isolated polypeptide (I) with retinoid acid
CC	(RA) pathway activity, comprising a polypeptide sequence of Perturbation
CC	(phenotypic probe) R3, Perturbation F802, Perturbation F820, their

biologically active modifications, or biologically active fragments. Also included are the polynucleotides encoding the perturbagens, a vector comprising the polynucleotide, preparing an RA pathway related polypeptide, a composition comprising the polypeptide, an antibody to the polypeptide, screening (M) putative RA-related therapeutics, by exposing a polypeptide/target interaction pair identified by a method utilizing the RA pathway polypeptide, to a number of agents, and recovering a subpopulation of disrupting agents which competitively displace the polypeptide from the target, where the disrupting agents are putative RA-related therapeutics. Also include are an isolated RA pathway polypeptide comprising PARI (a kinesin light chain-related protein) polypeptide and its encoding polynucleotide, a gene therapy vector comprising the RA pathway protein polypeptide or encoding or PARI and a host cell comprising the gene therapy vector. The RA pathway polypeptide is useful for identifying a cellular target that interacts with RA pathway-related polypeptide, by exposing the polypeptide in vitro to putative target molecules and identifying a polypeptide/target interaction pair, by detecting reporter expression, where the reporter expression is operatively linked to the formation of the interaction pair. (M) is a yeast two-hybrid assay. The polypeptide is also useful for treating an RA pathway-related condition e.g. lung cancer, Kaposi's sarcoma, breast cancer, pancreatic cancer, neuroblastoma, renal cancer, ovarian cancer, dermatitis, hyperkeratosis, eczema, Darier's disease, Reller's disease, psoriasis, acute promyelocytic leukemia (APL). The present sequence is a cDNA encoding partial perturbagen F820 and containing sequences derived from the vector (EGFP, inactive green fluorescent protein sequence).

Sequence 339 BP, 119 A, 57 C, 64 G, 99 T, 0 other;

Query Match 0.8%; Score 23; DB 24; Length 339;

Best Local Similarity 100.0%; Pred. No. 1.2;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AATCTTTATTTATCGATGTTA 23

332 AATCTTTATTTATCGATGTTA 310

RESULT 66

ABK87382/C

ID ABK87382 standard; cDNA, 352 BP.

AC ABK87382;

DT 24-SEP-2002 (first entry)

CDNA fragment for perturbagen R1.

Retinoic acid pathway; RA; ss; retinoid; lung cancer; Kaposi's sarcoma; breast cancer; pancreatic cancer; neuroblastoma; renal cancer; ovarian cancer; dermatitis; hyperkeratosis; eczema; Darier's disease; Reller's disease; psoriasis; gene therapy; R1; acute promyelocytic leukemia; APL; perturbagen; R3; F802; F820; yeast two-hybrid assay; PARI; kinesin light chain-related protein; human; EGFP; green fluorescent protein.

Homo sapiens.

Aequorea victoria.

Synthetic.

WO200240719-AZ.

23-MAY-2002.

17-NOV-2001; 2001WO-US44039.

17-NOV-2000; 2000US-249468P.

(DELT-) DELTAGEN PROTEOMICS INC.

Kamb CA, Richards BT, Karpilow J;

WPI; 2002-519386/55.

Polypeptide with retinoic acid pathway activity, especially of perturbagens R3, F802 and F820 for identifying a cellular target which interacts with the polypeptide and for therapeutic purposes

Example 4; Fig 15a; 131p; English.

The invention relates to an isolated polypeptide (I) with retinoic acid (RA) pathway activity, comprising a polypeptide sequence of Perturbagen (phenotypic probe) R3, Perturbagen F802, Perturbagen F820, their biologically active modifications, or biologically active fragments. Also included are the polynucleotides encoding the perturbagens, a vector comprising the polynucleotide, preparing an RA pathway related polypeptide, a composition comprising the polypeptide, an antibody to the polypeptide, screening (M) putative RA-related therapeutics, by exposing a polypeptide/target interaction pair identified by a method utilizing the RA pathway polypeptide, to a number of agents, and recovering a subpopulation of disrupting agents which competitively displace the polypeptide from the target, where the disrupting agents are putative RA-related therapeutics. Also include are an isolated RA pathway polypeptide comprising PARI (a kinesin light chain-related protein) polypeptide and its encoding polynucleotide, a gene therapy vector comprising the RA pathway protein polypeptide or encoding or PARI and a host cell comprising the gene therapy vector. The RA pathway polypeptide is useful for identifying a cellular target that interacts with RA pathway-related polypeptide, by exposing the polypeptide in vitro to putative target molecules and identifying a polypeptide/target interaction pair, by detecting reporter expression, where the reporter expression is operatively linked to the formation of the interaction pair. (M) is a yeast two-hybrid assay. The polypeptide is also useful for treating an RA pathway-related condition e.g. lung cancer, Kaposi's sarcoma, breast cancer, pancreatic cancer, neuroblastoma, renal cancer, ovarian cancer, dermatitis, hyperkeratosis, eczema, Darier's disease, Reller's disease, psoriasis, acute promyelocytic leukemia (APL). The present sequence is a cDNA fragment representing perturbagen R1 and containing sequences derived from the vector (EGFP, inactive green fluorescent protein sequence).

Sequence 352 BP, 77 A, 66 C, 92 G, 117 T, 0 other;

Query Match 0.8%; Score 23; DB 24; Length 352;

Best Local Similarity 100.0%; Pred. No. 1.2;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 AATCTTTATTTATCGATGTTA 23

346 AATCTTTATTTATCGATGTTA 324

RESULT 67

ABK87381/C

ID ABK87381 standard; cDNA, 527 BP.

AC ABK87381;

DT 24-SEP-2002 (first entry)

CDNA encoding partial perturbagen H241.

Retinoic acid pathway; RA; ss; retinoid; lung cancer; Kaposi's sarcoma; breast cancer; pancreatic cancer; neuroblastoma; renal cancer; ovarian cancer; dermatitis; hyperkeratosis; eczema; Darier's disease; Reller's disease; psoriasis; gene therapy; H241; acute promyelocytic leukemia; APL; perturbagen; R3; F802; F820; yeast two-hybrid assay; PARI; kinesin light chain-related protein; human; EGFP; green fluorescent protein.

Homo sapiens.

Aequorea victoria.

Synthetic.

Key Location/Qualifiers

FT CDS 44..83
FT /*tag= a
FT /product= "Partial perturbation H241"
FT /partial
FT /note= "No start codon shown"
XX
XX WO200240719-A2.
XX
XX 23-MAY-2002.
XX
XX 17-NOV-2001; 2001WO-US44039.
XX
XX 17-NOV-2000; 2000US-249468P.
XX
XX (DELTA-) DELTAGEN PROTEOMICS INC.
XX
XX Kamb CA, Richards BT, Karpilow J;
XX WPI; 2002-519386/55.
XX P-PSDB; AAU99183.
XX
XX Polypeptide with retinoic acid pathway activity, especially of
PT perturbation R3, F802 and F820 for identifying a cellular target which
PT interacts with the polypeptide and for therapeutic purposes
XX
XX Example 3; Fig 13b; 131pp; English.
XX
XX The invention relates to an isolated polypeptide (I) with retinoic acid
CC (RA) pathway activity, comprising a polypeptide sequence of Perturbagen
CC (phenotypic probe) R3, Perturbagen F802, Perturbagen F820, their
CC biologically active modifications, or biologically active fragments.
CC Also included are the polynucleotides encoding the perturbagens, a
CC vector comprising the polynucleotide, preparing an RA pathway related
CC polypeptide, a composition comprising the polypeptide, an antibody to the
CC polypeptide, screening (M) putative RA-related therapeutics, by exposing
CC a polypeptide/target interaction pair identified by a method utilizing
CC the RA pathway polypeptide, to a number of agents, and recovering a
CC subpopulation of disrupting agents which competitively displace the
CC polypeptide from the target, where the disrupting agents are putative RA-
CC related therapeutics. Also include are an isolated RA pathway polypeptide
CC comprising PRT1 (a kinesin light chain-related protein) polypeptide
CC and its encoding polynucleotide, a gene therapy vector comprising the RA
CC pathway protein polypeptide or encoding or PRT1 and a host cell
CC comprising the gene therapy vector. The RA pathway polypeptide is useful
CC for identifying a cellular target that interacts with RA pathway-related
CC polypeptide, by exposing the polypeptide in vitro to putative target
CC molecules and identifying a polypeptide/target interaction pair, by
CC detecting reporter expression, where the reporter expression is
CC operatively linked to the formation of the interaction pair. (M) is a
CC yeast two-hybrid assay. The polypeptide is also useful for treating
CC an RA pathway-related condition e.g. lung cancer, Kaposi's sarcoma,
CC breast cancer, pancreatic cancer, neuroblastoma, renal cancer,
CC ovarian cancer, dermatitis, hyperkeratosis, eczema, Darier's disease,
CC Reiter's disease, psoriasis, acute promyelocytic leukaemia (APL).
CC The present sequence is a cDNA encoding partial perturbation H241
CC and containing sequences derived from the vector (RGFP, inactive green
CC fluorescent protein sequence).
XX
XX Sequence 527 BP; 187 A; 81 C; 130 G; 129 T; 0 other;
SQ

Query Match 0.8%; Score 23; DB 24; Length 527;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTTATTTATGAGTTA 23
Db 520 AATCTTTTATTTATGAGTTA 498

RESULT 68
AAK63284
ID AAK63284 standard; cDNA; 632 BP.
XX

AC AAK63284;
XX
XX 06-NOV-2001 (first entry)
DT
XX Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:8344.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ss.
XX
XX Homo sapiens.
XX
XX WO200157182-A2.
XX
XX 09-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
XX 04-FEB-2000; 2000US-0180628.
XX 24-FEB-2000; 2000US-0184664.
XX 02-MAR-2000; 2000US-0186350.
XX 16-MAR-2000; 2000US-0189874.
XX 17-MAR-2000; 2000US-0190076.
XX 18-APR-2000; 2000US-0198123.
XX 19-MAY-2000; 2000US-0205515.
XX 07-JUN-2000; 2000US-0209467.
XX 28-JUN-2000; 2000US-0214886.
XX 30-JUN-2000; 2000US-0215135.
XX 07-JUL-2000; 2000US-0216647.
XX 07-JUL-2000; 2000US-0216880.
XX 11-JUL-2000; 2000US-0217487.
XX 11-JUL-2000; 2000US-0217496.
XX 14-JUL-2000; 2000US-0218290.
XX 26-JUL-2000; 2000US-0220963.
XX 26-JUL-2000; 2000US-0220964.
XX 14-AUG-2000; 2000US-0224518.
XX 14-AUG-2000; 2000US-0224519.
XX 14-AUG-2000; 2000US-0225213.
XX 14-AUG-2000; 2000US-0225214.
XX 14-AUG-2000; 2000US-0225266.
XX 14-AUG-2000; 2000US-0225267.
XX 14-AUG-2000; 2000US-0225268.
XX 14-AUG-2000; 2000US-0225270.
XX 14-AUG-2000; 2000US-0225447.
XX 14-AUG-2000; 2000US-0225757.
XX 14-AUG-2000; 2000US-0225758.
XX 14-AUG-2000; 2000US-0225759.
XX 14-AUG-2000; 2000US-0226279.
XX 18-AUG-2000; 2000US-0226881.
XX 22-AUG-2000; 2000US-0226886.
XX 22-AUG-2000; 2000US-0227182.
XX 23-AUG-2000; 2000US-0227009.
XX 30-AUG-2000; 2000US-0228924.
XX 01-SEP-2000; 2000US-0228287.
XX 01-SEP-2000; 2000US-0228343.
XX 01-SEP-2000; 2000US-0228344.
XX 01-SEP-2000; 2000US-0228345.
XX 05-SEP-2000; 2000US-0228309.
XX 05-SEP-2000; 2000US-0228513.
XX 06-SEP-2000; 2000US-0230437.
XX 06-SEP-2000; 2000US-0230438.
XX 08-SEP-2000; 2000US-0231243.
XX 08-SEP-2000; 2000US-0231244.
XX 08-SEP-2000; 2000US-0231244.
XX 08-SEP-2000; 2000US-0231413.
XX 08-SEP-2000; 2000US-0231414.
XX 08-SEP-2000; 2000US-0232080.
XX 08-SEP-2000; 2000US-0232081.
XX 12-SEP-2000; 2000US-0231968.
XX 14-SEP-2000; 2000US-0232397.
XX 14-SEP-2000; 2000US-0232398.
XX 14-SEP-2000; 2000US-0232399.
XX 14-SEP-2000; 2000US-0232400.

PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.
 PR 26-SEP-2000; 2000US-0235484.
 PR 27-SEP-2000; 2000US-0235834.
 PR 27-SEP-2000; 2000US-0235836.
 PR 29-SEP-2000; 2000US-0235327.
 PR 29-SEP-2000; 2000US-0235367.
 PR 29-SEP-2000; 2000US-0235368.
 PR 29-SEP-2000; 2000US-0235369.
 PR 29-SEP-2000; 2000US-0235370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 02-OCT-2000; 2000US-0237040.
 PR 13-OCT-2000; 2000US-0239335.
 PR 13-OCT-2000; 2000US-0239337.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0241826.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0255719.
 PR 06-DEC-2000; 2000US-0251479.

PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX
 PI Rosen CA, Baraeth SC, Ruben SM;
 DR WPI; 2001-483426/52.
 DR P-PSDB; AAM90503.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 FT
 XX
 XX Claim 1; SEQ ID NO 8344; 3071bp + Sequence Listing; English.
 XX
 PS AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting the
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 632 BP; 202 A; 131 C; 124 G; 170 T; 5 other;
 Query Match 0.8%; Score 23; DB 22; Length 632;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 57 GTGACCCACGCGTCCGGAGTA 79
 DB 72 GTGACCCACGCGTCCGGAGTA 94
 RESULT 69
 ID ABR87378 standard; cDNA; 661 BP.
 XX
 AC ABR87378;
 XX
 DT 24-SEP-2002 (first entry)
 XX
 DB cDNA encoding partial perturbation F797.
 XX
 KW Retinoic acid pathway; RA; ss; retinoid; lung cancer;
 KW Kaposi's sarcoma; breast cancer; pancreatic cancer; neuroblastoma;
 KW renal cancer; ovarian cancer; dermatitis; hyperkeratosis; eczema;
 KW Darier's disease; Reiter's disease; psoriasis; gene therapy; F797;
 KW acute promyelocytic leukaemia; APL; Perturbation; R3; F802; F820;
 KW yeast two-hybrid assay; PAVI; kinesin light chain-related protein;
 KW human; EGFP; green fluorescent protein.
 XX
 OS Homo sapiens.
 OS Aequorea victoria.
 OS Synthetic.
 XX
 Key Location/Qualifiers

FT CDS 44.112
 FT /*tag= a
 FT /product= "Partial perturbation F797"
 FT /partial
 FT /note= "No start codon shown"
 FT /transl_except= (pos:95..97,aa:leu)
 PN WO200240719-A2.
 XX
 XX 23-MAY-2002.
 PD
 XX
 XX 17-NOV-2001; 2001WO-US44039.
 XX
 XX 17-NOV-2000; 2000US-249468P.
 XX
 XX (DELTA-) DELTAGEN PROTEOMICS INC.
 XX
 XX Kamb CA, Richards BT, Karpilow J;
 PI
 XX WPI; 2002-519386/55.
 DR P-PSDB; AAU99180.
 XX
 XX Polypeptide with retinoic acid pathway activity, especially of
 PT perturbation R3, F802 and F820 for identifying a cellular target which
 PT interacts with the polypeptide and for therapeutic purposes -
 XX
 XX Example 3; Fig 13a; 131pp; English.

CC The invention relates to an isolated polypeptide (I) with retinoic acid
 CC (RA) pathway activity, comprising a polypeptide sequence of Perturbation
 CC (phenotypic probe) R3, Perturbation F802, Perturbation F820, their
 CC biologically active modifications, or biologically active fragments.
 CC Also included are the polynucleotides encoding the perturbagens, a
 CC vector comprising the polynucleotide, preparing an RA pathway related
 CC polypeptide, a composition comprising the polypeptide, an antibody to the
 CC polypeptide, screening (M) putative RA-related therapeutics, by exposing
 CC a polypeptide/target interaction pair identified by a method utilizing
 CC the RA pathway polypeptide, to a number of agents, and recovering a
 CC subpopulation of disrupting agents which competitively displace the
 CC polypeptide from the target, where the disrupting agents are putative RA-
 CC related therapeutics. Also include are an isolated RA pathway polypeptide
 CC comprising P81 (a kinesin light chain-related protein) polypeptide
 CC and its encoding polynucleotide, a gene therapy vector comprising the RA
 CC pathway protein polypeptide or encoding or P81 and a host cell
 CC comprising the gene therapy vector. The RA pathway polypeptide is useful
 CC for identifying a cellular target that interacts with RA pathway-related
 CC polypeptide, by exposing the polypeptide in vitro to putative target
 CC molecules and identifying a polypeptide/target interaction pair, by
 CC detecting reporter expression, where the reporter expression is
 CC operatively linked to the formation of the interaction pair. (M) is a
 CC yeast two-hybrid assay. The polypeptide is also useful for treating
 CC an RA pathway-related condition e.g. lung cancer, Kaposi's sarcoma,
 CC breast cancer, pancreatic cancer, neuroblastoma, renal cancer,
 CC ovarian cancer, dermatitis, hyperkeratosis, eczema, Darier's disease,
 CC Reiter's disease, psoriasis, acute promyelocytic leukemia (APL).
 CC The present sequence is a cDNA encoding partial perturbation F797
 CC and containing sequences derived from the vector (BGFP, inactive green
 CC fluorescent protein sequence).
 XX
 SQ Sequence 661 BP; 248 A; 87 C; 156 G; 170 T; 0 other;

Query Match 0.8%; Score 23; DB 24; Length 661;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AATCTTTATTATTCGATGTTA 23
 |||
 DB 654 AATCTTTATTATTCGATGTTA 632

RESULT 70
 AAH36207
 ID AAH36207 standard; cDNA; 115 BP.

XX
 AC AAH36207;
 XX
 XX 03-SEP-2001 (first entry)
 DT
 XX Human colon cancer antigen encoding cDNA SEQ ID NO:3289.
 DE
 XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
 KW colorectal carcinoma; SB.
 XX
 XX Homo sapiens.
 OS
 XX WO200122920-A2.
 PN
 XX 05-APR-2001.
 PD
 XX
 XX 28-SEP-2000; 2000WO-US26524.
 PF
 XX 29-SEP-1999; 99US-0157137.
 PR 03-NOV-1999; 99US-0163280.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Ruben SM, Barash SC, Birse CE, Rosen CA;
 PI
 XX WPI; 2001-235357/24.
 DR P-PSDB; AAG76802.
 XX
 XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
 PT useful for preventing, diagnosing and/or treating colorectal cancers -
 XX
 XX Claim 1; Page 5085-5086; 9803pp; English.

CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
 CC cancer-associated nucleic acid molecules (N) and proteins (P), where
 CC the proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene
 CC therapy and vaccine production. N and P may be used in the prevention,
 CC diagnosis and treatment of diseases associated with inappropriate P
 CC expression. For example, N and P may be used to treat disorders
 CC associated with decreased expression by rectifying mutations or deletions
 CC in a patient's genome that affect the activity of P by expressing
 CC inactive proteins or to supplement the patient's own production of P.
 CC Additionally, N may be used to produce the colon cancer-associated Ps,
 CC by inserting the nucleic acids into a host cell and culturing the cell
 CC to express the proteins. N and P can be used in the prevention, diagnosis
 CC and treatment of colorectal carcinomas and cancers. AAH37196 to AAH37204
 CC and AAG77789 represent sequences used in the exemplification of the
 CC present invention.
 CC N.B.: Pages 666 to 682 and page 7053 of the sequence listing were
 CC missing at time of publication, meaning no sequences are present for
 CC SEQ ID NO:1027 to 1052, 7921 and 7922.

SQ Sequence 115 BP; 20 A; 33 C; 34 G; 21 T; 7 other;

Query Match 0.8%; Score 22; DB 22; Length 115;
 Best Local Similarity 100.0%; Pred. No. 3.7;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTCGACCCACCGCTCCGGAGT 78
 |||
 DB 61 GTCGACCCACCGCTCCGGAGT 82

RESULT 71
 AAH61215
 ID AAH61215 standard; cDNA; 370 BP.
 XX
 AC AAH61215;
 AC
 XX 06-NOV-2001 (first entry)
 DT
 XX Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:6275.

Human; immune; hematopoietic; immune/haematopoietic antigen; cancer;
cytostatic; gene therapy; vaccine; metastasis; ss.
Homo sapiens.
WO200157182-A2.
09-AUG-2001.
17-JAN-2001; 2001WO-US01354.
31-JAN-2000; 2000US-0179065.
04-FEB-2000; 2000US-0180628.
24-FEB-2000; 2000US-0184664.
02-MAR-2000; 2000US-0186350.
16-MAR-2000; 2000US-0189874.
17-MAR-2000; 2000US-0190076.
18-APR-2000; 2000US-0198123.
19-MAY-2000; 2000US-0205515.
07-JUN-2000; 2000US-0209467.
28-JUN-2000; 2000US-0214886.
30-JUN-2000; 2000US-0215135.
07-JUL-2000; 2000US-0216647.
07-JUL-2000; 2000US-0216880.
11-JUL-2000; 2000US-0217487.
11-JUL-2000; 2000US-0217496.
14-JUL-2000; 2000US-0218290.
26-JUL-2000; 2000US-0220964.
26-JUL-2000; 2000US-0220964.
14-AUG-2000; 2000US-0224518.
14-AUG-2000; 2000US-0224519.
14-AUG-2000; 2000US-0225213.
14-AUG-2000; 2000US-0225214.
14-AUG-2000; 2000US-0225265.
14-AUG-2000; 2000US-0225267.
14-AUG-2000; 2000US-0225268.
14-AUG-2000; 2000US-0225270.
14-AUG-2000; 2000US-0225447.
14-AUG-2000; 2000US-0225757.
14-AUG-2000; 2000US-0225758.
14-AUG-2000; 2000US-0225759.
18-AUG-2000; 2000US-0226279.
22-AUG-2000; 2000US-0226681.
22-AUG-2000; 2000US-0226868.
22-AUG-2000; 2000US-0227182.
23-AUG-2000; 2000US-0227009.
30-AUG-2000; 2000US-0228924.
01-SEP-2000; 2000US-0229287.
01-SEP-2000; 2000US-0229343.
01-SEP-2000; 2000US-0229344.
01-SEP-2000; 2000US-0229345.
05-SEP-2000; 2000US-0229509.
05-SEP-2000; 2000US-0229513.
06-SEP-2000; 2000US-0230437.
06-SEP-2000; 2000US-0230438.
08-SEP-2000; 2000US-0231243.
08-SEP-2000; 2000US-0231244.
08-SEP-2000; 2000US-0231244.
08-SEP-2000; 2000US-0231414.
08-SEP-2000; 2000US-0231414.
08-SEP-2000; 2000US-0232080.
08-SEP-2000; 2000US-0232081.
12-SEP-2000; 2000US-0231968.
14-SEP-2000; 2000US-0232397.
14-SEP-2000; 2000US-0232398.
14-SEP-2000; 2000US-0232399.
14-SEP-2000; 2000US-0232400.
14-SEP-2000; 2000US-0232401.
14-SEP-2000; 2000US-0233063.
14-SEP-2000; 2000US-0233064.
14-SEP-2000; 2000US-0233065.
21-SEP-2000; 2000US-0234223.
21-SEP-2000; 2000US-0234274.
25-SEP-2000; 2000US-0234997.
25-SEP-2000; 2000US-0234998.
26-SEP-2000; 2000US-0235484.
27-SEP-2000; 2000US-0235834.
27-SEP-2000; 2000US-0235836.
29-SEP-2000; 2000US-0236327.
29-SEP-2000; 2000US-0236367.
29-SEP-2000; 2000US-0236368.
29-SEP-2000; 2000US-0236369.
29-SEP-2000; 2000US-0236370.
02-OCT-2000; 2000US-0236802.
02-OCT-2000; 2000US-0237037.
02-OCT-2000; 2000US-0237038.
02-OCT-2000; 2000US-0237039.
02-OCT-2000; 2000US-0237040.
13-OCT-2000; 2000US-0239935.
13-OCT-2000; 2000US-0239937.
20-OCT-2000; 2000US-0240960.
20-OCT-2000; 2000US-0241221.
20-OCT-2000; 2000US-0241785.
20-OCT-2000; 2000US-0241786.
20-OCT-2000; 2000US-0241787.
20-OCT-2000; 2000US-0241808.
20-OCT-2000; 2000US-0241809.
01-NOV-2000; 2000US-0241826.
08-NOV-2000; 2000US-0244617.
08-NOV-2000; 2000US-0246474.
08-NOV-2000; 2000US-0246475.
08-NOV-2000; 2000US-0246476.
08-NOV-2000; 2000US-0246477.
08-NOV-2000; 2000US-0246478.
08-NOV-2000; 2000US-0246523.
08-NOV-2000; 2000US-0246524.
08-NOV-2000; 2000US-0246525.
08-NOV-2000; 2000US-0246526.
08-NOV-2000; 2000US-0246527.
08-NOV-2000; 2000US-0246528.
08-NOV-2000; 2000US-0246532.
08-NOV-2000; 2000US-0246609.
08-NOV-2000; 2000US-0246610.
08-NOV-2000; 2000US-0246611.
08-NOV-2000; 2000US-0246613.
17-NOV-2000; 2000US-0249207.
17-NOV-2000; 2000US-0249208.
17-NOV-2000; 2000US-0249209.
17-NOV-2000; 2000US-0249210.
17-NOV-2000; 2000US-0249211.
17-NOV-2000; 2000US-0249212.
17-NOV-2000; 2000US-0249213.
17-NOV-2000; 2000US-0249214.
17-NOV-2000; 2000US-0249215.
17-NOV-2000; 2000US-0249216.
17-NOV-2000; 2000US-0249217.
17-NOV-2000; 2000US-0249218.
17-NOV-2000; 2000US-0249244.
17-NOV-2000; 2000US-0249245.
17-NOV-2000; 2000US-0249245.
17-NOV-2000; 2000US-0249264.
17-NOV-2000; 2000US-0249265.
17-NOV-2000; 2000US-0249297.
17-NOV-2000; 2000US-0249299.
01-DEC-2000; 2000US-0250160.
01-DEC-2000; 2000US-0250391.
05-DEC-2000; 2000US-0251030.
05-DEC-2000; 2000US-0251988.
05-DEC-2000; 2000US-0251988.
06-DEC-2000; 2000US-0251479.
08-DEC-2000; 2000US-0251856.
08-DEC-2000; 2000US-0251858.
08-DEC-2000; 2000US-0251869.
08-DEC-2000; 2000US-0251989.
08-DEC-2000; 2000US-0251990.

```

PR 11-DEC-2000; 2000US-0254097.
XX 05-JAN-2001; 2001US-0259678.
XX (HUMA-) HUMAN GENOME SCI INC.
XX PI Rosen CA, Barash SC, Ruben SM;
XX DR WPI; 2001-483426/52.
XX P-PSDB; AAM88434.
XX
PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
PT useful for preventing, diagnosing and/or treating cancers and
PT metastasis -
XX
PS Claim 1; SEQ ID NO 6275; 3071pp + Sequence Listing; English.
XX
CC AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
CC activity, and can be used in gene therapy and vaccine production. (I)
CC proteins and polynucleotides may be used in the prevention, diagnosis and
CC treatment of diseases associated with inappropriate (I) expression. For
CC example, they may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of (I) by expressing inactive proteins or to
CC supplement the patient's own production of (I). Additionally, (I)
CC polynucleotides may be used to produce the secreted (I), by inserting the
CC the nucleic acids into a host cell and culturing the cell to express the
CC protein. (I) proteins and polynucleotides may be used to prevent,
CC diagnose and treat immune/hematopoietic-related diseases, especially
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703
CC to AAK67694 represent human immune/haematopoietic antigen genomic
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
CC represent sequences used in the exemplification of the present invention.
XX
SQ Sequence 370 BP; 62 A; 90 C; 146 G; 67 T; 5 other;

```

Query Match 0.8%; Score 22; DB 22; Length 370;
 Best Local Similarity 100.0%; Pred. No. 3.6;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 57 GTGACCCACGCGTCGGGAGT 78
DB 3 GTGACCCACGCGTCGGGAGT 24

```

RESULT 72
 AA192960
 ID AA192960 standard; cDNA; 471 BP.
 XX
 AC AA192960;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 13020.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorders; arthritis; inflammation; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200164835-A2.
 XX
 PD 07-SEP-2001.
 XX
 PF 26-FEB-2001; 2001WO-US04927.
 XX
 PR 28-FEB-2000; 2000US-0515126.
 XX
 PR 18-MAY-2000; 2000US-0577409.
 XX
 PA (HYSE-) HYSEQ INC.

```

PI Tang YT, Liu C, Dirmagac RT;
XX
XX WPI; 2001-514838/56.
XX DR P-PSDB; AAO13029.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
PS Claim 1; SEQ ID NO 13020; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AA179941-AA193841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 471 BP; 80 A; 151 C; 112 G; 123 T; 5 other;

```

Query Match 0.8%; Score 22; DB 22; Length 471;
 Best Local Similarity 100.0%; Pred. No. 3.6;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 57 GTGACCCACGCGTCGGGAGT 78
DB 43 GTGACCCACGCGTCGGGAGT 64

```

RESULT 73
 ABS77002
 ID ABS77002 standard; cDNA; 711 BP.
 XX
 AC ABS77002;
 XX
 DT 12-DEC-2002 (first entry)
 XX
 DE Frog embryonic gene sequence Q925394.
 XX
 KW Frog; ss; embryonic development; developmental disorder;
 KW microarray; cell differentiation.
 XX
 OS Xenopus laevis.
 XX
 PN US2002081610-A1.
 XX
 PD 27-JUN-2002.
 XX
 PF 23-JUL-2001; 2001US-0910943.
 XX
 PR 21-JUL-2000; 2000US-219658P.
 XX
 PA (UYRQ) UNIV ROCKEFELLER.
 XX
 PA Hemmati-Briavanlou A, Altmann CR;
 XX
 PI WPI; 2002-626534/67.
 XX
 DR 26-FEB-2001; 2001WO-US04927.
 XX
 PF Nucleic acid array containing Xenopus embryonic nucleic acids is useful
 PF to identify genes involved in embryonic development, to identify
 PF different types of embryonic cells, and to diagnose developmental
 PF disorders -
 XX
 PS Claim 1; Page 272; 823pp; English.

CC The invention relates to a nucleic acid array, where each coordinate
 CC contains a single nucleic acid species having one of 770 nucleotide
 CC sequences (appearing as ABS76747-ABS77516) a of a Xenopus embryonic gene
 CC product, or its complement or hybridisable fragment of not less than
 CC 20 contiguous nucleotides of one of those sequences. Also included are
 CC detecting differential expression of embryonic genes, comprising:
 CC (a) contacting a nucleic acid array comprising genes expressed in
 CC embryonic but not mature cells with nucleic acids from sample and
 CC control cells; and (b) detecting differential hybridisation of nucleic
 CC acids from the sample cells relative to the control cells; and detecting
 CC defects in development, comprising: (a) contacting nucleic acids from
 CC test cells undergoing development with a nucleic acid array of gene
 CC products known to play a fundamental role in the development process; and
 CC (b) detecting a difference in expression of a fundamental gene in the
 CC sample cells relative to a standard. The invention is useful to identify
 CC genes involved in embryonic development and related processes such as
 CC cell differentiation. This would be useful for diagnosing developmental
 CC disorders and for identifying different types of embryonic cells.
 CC The present sequence is one of the 770 Xenopus embryonic cDNA sequences.

SO Sequence 711 BP; 219 A; 173 C; 170 G; 148 T; 1 other;
 Query Match 0.8%; Score 22; DB 24; Length 711;
 Best Local Similarity 100.0%; Pred. No. 3.6;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAGT 78
 |||||
 DB 48 GTGACCCACGCGTCCGGAGT 69

RESULT 74
 ABS77128
 ID ABS77128 standard; cDNA; 748 BP.

AC ABS77128;
 DT 12-DEC-2002 (first entry)

DE Frog embryonic gene sequence Q9925519.

KW Frog; ss; embryonic development; developmental disorder;
 KW microarray; cell differentiation.

OS Xenopus laevis.

PN US2002081610-A1.

PD 27-JUN-2002.

PF 23-JUL-2001; 2001US-0910943.

PR 21-JUL-2000; 2000US-219658P.

PA (UYRQ) UNIV ROCKEFELLER.

PI Hemmati-Briavanlou A, Altman CR;

DR WPI; 2002-626534/67.

PT Nucleic acid array containing Xenopus embryonic nucleic acids is useful
 PT to identify genes involved in embryonic development, to identify
 PT different types of embryonic cells, and to diagnose developmental
 PT disorders -

PS Claim 1; Page 398; 823bp; English.

CC The invention relates to a nucleic acid array, where each coordinate
 CC contains a single nucleic acid species having one of 770 nucleotide
 CC sequences (appearing as ABS76747-ABS77516) a of a Xenopus embryonic gene
 CC product, or its complement or hybridisable fragment of not less than
 CC 20 contiguous nucleotides of one of those sequences. Also included are
 CC detecting differential expression of embryonic genes, comprising:

CC (a) contacting a nucleic acid array comprising genes expressed in
 CC embryonic but not mature cells with nucleic acids from sample and
 CC control cells; and (b) detecting differential hybridisation of nucleic
 CC acids from the sample cells relative to the control cells; and detecting
 CC defects in development, comprising: (a) contacting nucleic acids from
 CC test cells undergoing development with a nucleic acid array of gene
 CC products known to play a fundamental role in the development process; and
 CC (b) detecting a difference in expression of a fundamental gene in the
 CC sample cells relative to a standard. The invention is useful to identify
 CC genes involved in embryonic development and related processes such as
 CC cell differentiation. This would be useful for diagnosing developmental
 CC disorders and for identifying different types of embryonic cells.
 CC The present sequence is one of the 770 Xenopus embryonic cDNA sequences.

SO Sequence 748 BP; 237 A; 122 C; 137 G; 216 T; 36 other;
 Query Match 0.8%; Score 22; DB 24; Length 748;
 Best Local Similarity 100.0%; Pred. No. 3.6;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAGT 78
 |||||
 DB 49 GTGACCCACGCGTCCGGAGT 70

RESULT 75
 AAA87682
 ID AAA87682 standard; cDNA; 2112 BP.

AC AAA87682;

DT 04-DEC-2000 (first entry)

DE Human secreted protein gene 17 SEQ ID NO:27.

KW Human; secreted protein; immunosuppressive; immunostimulant; nootropic;
 KW antiinflammatory; cardiant; vulnerary; anticulcer; anticonvulsant;
 KW antiparkinsonian; neuroprotective; antiviral; antibacterial; cytostatic;
 KW antiparasitic; thrombolytic; anticoagulant; antiarteriosclerotic;
 KW gene therapy; vaccine; chemotaxis-modulator; angiogenesis-modulator;
 KW cancer; immune system disorder; hyperproliferative disorder; infection;
 KW cardiovascular disorder; neurological disease; wound healing; ss.

OS Homo sapiens.

PN WO200043495-A2.

PD 27-JUL-2000.

PF 18-JAN-2000; 2000WO-US00903.

PR 19-JAN-1999; 99US-0116330.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Ruben SM, Ebner R, Young PE, Ni J, Moore PA;

DR Komatsu G, Birse CB;

DR WPI; 2000-499225/44.

DR P-PSDB; AAB25681.

PT New isolated polynucleotide encoding a secreted protein useful for
 PT preventing, treating or ameliorating a medical condition -

PS Claim 1; Page 388; 451bp; English.

CC The polynucleotide sequences given in AAA87666 to AAA87708 encodes the
 CC human secreted proteins given in AAB25665 to AAB25755. Human secreted
 CC proteins have activities based on the tissues and cells the genes are
 CC expressed in. Examples of activities include: immunosuppressive;
 CC immunostimulant; antiinflammatory; cardiant; vulnerary; anticulcer;
 CC nootropic; antiviral; anticonvulsant; antiparkinsonian; neuroprotective;
 CC antibacterial; antiparasitic; thrombolytic; anticoagulant;

CC antiarteriosclerotic and cyostatic. The secreted proteins and their
CC polynucleotides can be used in gene therapy and as vaccines,
CC chemotaxis-modulators and angiogenesis- modulators. The human secreted
CC proteins and polynucleotides can be used for diagnosing (the
CC susceptibility to) a pathological condition by determining the presence
CC or absence of a mutation in the polynucleotide or determining the
CC presence or amount of expression of the protein. The polynucleotides and
CC proteins can also be used in the treatment and diagnosis of cancer,
CC diseases of the immune system, hyperproliferative disorders,
CC cardiovascular disorders and neurological disease. They can also be used
CC to promote wound healing and to fight infection. AA87657 to AA87665 and
CC AA85664 represent sequences used in the exemplification of the present
CC invention.
XX

SQ Sequence 2112 BP; 512 A; 450 C; 480 G; 657 T; 13 other;

Query Match 0.8%; Score 22; DB 21; Length 2112;

Best Local Similarity 100.0%; Pred. No. 3.5;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACGCGTCCGGAGT 78
|||||

DB 71 GTGACCCACGCGTCCGGAGT 92
|||||

Search completed: July 31, 2003, 19:48:34
Job time : 685 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 22:28:41 ; Search time 596 Seconds
(without alignments)
9376.961 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709
Sequence: 1 aactcttattatcagatg.....aggctttttcccaatacc 2709

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 1439767 seqs, 1031500376 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCCT_NEW_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2709	100.0	2709	US-09-774-490-1	Sequence 1, Appli
2	2508	92.6	2530	US-10-097-340-283	Sequence 283, App
3	2508	92.6	2530	US-10-263-538-9	Sequence 9, Appli
4	405	15.0	456	US-09-864-761-30992	Sequence 30992, A
5	303	11.2	354	US-09-922-217-736	Sequence 736, App
6	303	11.2	354	US-09-833-263-736	Sequence 736, App
7	303	11.2	354	US-10-025-380-736	Sequence 736, App
8	292	10.8	496	US-09-864-761-14444	Sequence 14444, A
9	160	5.9	172	US-09-864-761-27943	Sequence 27943, A
10	160	5.9	172	US-09-864-761-11363	Sequence 11363, A
11	144	5.3	446	US-09-864-761-11810	Sequence 11810, A
12	123	4.5	123	US-09-864-761-28381	Sequence 28381, A
13	122	4.5	456	US-09-864-761-6056	Sequence 6056, App
14	101	3.7	101	US-09-864-761-22893	Sequence 22893, A
15	71	2.6	2137	US-09-774-490-6	Sequence 6, Appli
16	67	2.5	520	US-10-027-632-238542	Sequence 238542, A
17	67	2.5	520	US-10-027-632-238543	Sequence 238543, A
18	67	2.5	520	US-09-897-006-10	Sequence 10, Appli
19	31	1.1	4661	US-09-897-511A-10	Sequence 10, Appli
20	31	1.1	7160	US-08-786-531B-5	Sequence 5, Appli
21	31	1.1	7235	US-08-786-531B-6	Sequence 6, Appli
22	26	1.0	1313	US-09-728-422-1	Sequence 1, Appli
23	25	0.9	1400	US-09-774-490-5	Sequence 5, Appli
24	22	0.8	115	US-10-106-698-32299	Sequence 3299, App
25	22	0.8	711	US-09-910-943-329	Sequence 236, App
26	22	0.8	748	US-09-910-943-361	Sequence 361, App
27	22	0.8	5732	US-09-897-006-6	Sequence 6, Appli
28	22	0.8	5732	US-09-897-511A-6	Sequence 6, Appli
29	22	0.8	7352	US-08-786-531B-4	Sequence 4, Appli
30	22	0.8	7353	US-08-786-531B-1	Sequence 1, Appli
31	22	0.8	9183	US-09-897-006-7	Sequence 7, Appli
32	22	0.8	9183	US-09-897-511A-7	Sequence 7, Appli
33	21	0.8	332	US-09-924-035A-528	Sequence 528, App
34	21	0.8	442	US-09-925-300-828	Sequence 828, App
35	21	0.8	457	US-09-960-352-4373	Sequence 4373, App
36	21	0.8	690	US-10-027-632-235696	Sequence 235696, A
37	21	0.8	717	US-10-106-698-1858	Sequence 1858, App
38	21	0.8	725	US-09-910-943-67	Sequence 67, Appli
39	21	0.8	730	US-09-910-943-448	Sequence 448, App
40	21	0.8	745	US-09-910-943-674	Sequence 674, App
41	21	0.8	749	US-09-910-943-359	Sequence 359, App
42	21	0.8	749	US-09-910-943-428	Sequence 428, App
43	21	0.8	750	US-09-910-943-365	Sequence 365, App
44	21	0.8	750	US-09-910-943-383	Sequence 504, App
45	21	0.8	781	US-09-910-943-383	Sequence 383, App
46	21	0.8	784	US-09-910-943-609	Sequence 609, App
47	21	0.8	883	US-10-198-846-14063	Sequence 10663, A
48	21	0.8	1181	US-10-037-270-33	Sequence 33, Appli
49	21	0.8	1280	US-10-106-698-1342	Sequence 1342, App
50	21	0.8	1385	US-09-925-776-1	Sequence 1, Appli
51	21	0.8	1394	US-10-119-926-44	Sequence 44, Appli
52	21	0.8	1572	US-10-037-270-664	Sequence 664, App
53	21	0.8	1808	US-09-838-573-1	Sequence 1, Appli
54	21	0.8	1808	US-10-160-501-1	Sequence 1, Appli
55	21	0.8	2570	US-09-925-302-192	Sequence 192, App
56	21	0.8	2810	US-09-796-753-89	Sequence 89, Appli
57	21	0.8	2815	US-09-782-960-67	Sequence 67, Appli
58	21	0.8	4207	US-09-897-006-4	Sequence 4, Appli
59	21	0.8	4207	US-09-897-511A-4	Sequence 4, Appli
60	21	0.8	4210	US-09-897-006-5	Sequence 5, Appli
61	21	0.8	4210	US-09-897-511A-5	Sequence 5, Appli
62	21	0.8	6255	US-09-897-006-13	Sequence 13, Appli
63	21	0.8	6255	US-09-897-511A-13	Sequence 13, Appli
64	21	0.8	9511	US-09-897-006-34	Sequence 34, Appli
65	21	0.8	9511	US-09-897-511A-34	Sequence 34, Appli
66	20	0.7	100	US-10-106-698-4090	Sequence 4090, App
67	20	0.7	122	US-09-764-891-522	Sequence 522, App
68	20	0.7	142	US-10-106-698-4018	Sequence 4018, App
69	20	0.7	197	US-10-106-698-2365	Sequence 2365, App
70	20	0.7	327	US-09-918-995-8601	Sequence 8601, App
71	20	0.7	355	US-09-986-480-113	Sequence 113, App
72	20	0.7	367	US-10-198-846-10302	Sequence 10302, A
73	20	0.7	402	US-09-925-300-802	Sequence 802, App
74	20	0.7	405	US-10-106-698-2172	Sequence 2172, App
75	20	0.7	430	US-09-960-352-13318	Sequence 13318, A
76	20	0.7	435	US-09-764-869-326	Sequence 326, App
77	20	0.7	435	US-10-091-504-326	Sequence 326, App
78	20	0.7	445	US-10-106-698-229	Sequence 229, App
79	20	0.7	453	US-09-918-995-7946	Sequence 7946, App
80	20	0.7	464	US-09-918-995-8118	Sequence 8118, App
81	20	0.7	474	US-10-102-806-41	Sequence 41, Appli
82	20	0.7	475	US-09-918-995-6137	Sequence 6137, App
83	20	0.7	481	US-09-918-995-7696	Sequence 7696, App
84	20	0.7	489	US-09-908-711-49	Sequence 49, Appli
85	20	0.7	481	US-09-764-891-567	Sequence 567, App
86	20	0.7	486	US-09-925-301-232	Sequence 232, App
87	20	0.7	491	US-09-918-995-7344	Sequence 7344, App
88	20	0.7	533	US-09-924-035A-540	Sequence 540, App
89	20	0.7	556	US-09-908-711-51	Sequence 51, Appli

90	20	0.7	556	11	US-09-986-480-61	Sequence 64, App
91	20	0.7	556	11	US-09-764-691-570	Sequence 570, App
92	20	0.7	557	14	US-10-102-806-198	Sequence 198, App
93	20	0.7	558	13	US-10-037-632-13709	Sequence 13709, App
94	20	0.7	606	10	US-09-764-877-115	Sequence 115, App
95	20	0.7	612	11	US-09-986-480-615	Sequence 55, App
96	20	0.7	632	10	US-09-764-684-1729	Sequence 429, App
97	20	0.7	633	14	US-10-106-638-1763	Sequence 1763, App
98	20	0.7	636	9	US-09-910-943-288	Sequence 288, App
99	20	0.7	702	9	US-09-910-943-240	Sequence 240, App
100	20	0.7	717	9	US-09-910-943-123	Sequence 123, App

ALIGNMENTS

```

RESULT 1
US-09-774-490-1
; Sequence 1, Application US/09774490
; Patent No. US2001003432A1
; GENERAL INFORMATION:
; APPLICANT: jin, Shengfang
; TITLE OF INVENTION: RESISTANCE SEQUENCES AND USES THEREOF
; FILE REFERENCE: 07334-138001
; CURRENT APPLICATION NUMBER: US/09/774,490
; CURRENT FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: US 60/179,191
; PRIOR FILING DATE: 2000-01-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2709
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-774-490-1

```

Query Match	100.0%;	Score 2709;	DB 9;	Length 2709;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2709; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0

Qy	1	AATCTTTATTTATTCGATGTTAAACAGCTTGATTCATCCACGTCGAGGGGTGCG	60
Db	1	AATCTTTATTTATTCGATGTTAAACAGCTTGATTCATCCACGTCGAGGGGTGCG	60
Qy	61	ACCCAGCGCTGCGGAGTAGTGTGAGCTGCGCTGTTCCATTGTCAAGCATATT	120
Db	61	ACCCAGCGCTGCGGAGTAGTGTGAGCTGCGCTGTTCCATTGTCAAGCATATT	120
Qy	121	CCAGATTGTTTGAACCTCTCTGGCGCGCAACATACAGAAAGAAAGACTTAAACAGCAAAAG	180
Db	121	CCAGATTGTTTGAACCTCTCTGGCGCGCAACATACAGAAAGAAAGACTTAAACAGCAAAAG	180
Qy	181	GACCTACAGCTGTGACAGATGGGCTGGTTAACTAGATTTGTCTGCTTTTCTGGGAGT	240
Db	181	GACCTACAGCTGTGACAGATGGGCTGGTTAACTAGATTTGTCTGCTTTTCTGGGAGT	240
Qy	241	ATTACTTACGCAAGAGCAAACTATAGAAATGGGAAAGCAATGTGCCAAGCTGAATT	300
Db	241	ATTACTTACGCAAGAGCAAACTATAGAAATGGGAAAGCAATGTGCCAAGCTGAATT	300
Qy	301	ATCTTCAAAAGAAATGTGGAATCCAAACAATGTGATCACTTTCAAATGGCTTGGCCAAAG	360
Db	301	ATCTTCAAAAGAAATGTGGAATCCAAACAATGTGATCACTTTCAAATGGCTTGGCCAAAG	360
Qy	361	CTCCAGTTATCAACCTCTCTTTGGATAGAGAAAGAGTAGGCTGTATGTTTGGAGCAA	420
Db	361	CTCCAGTTATCAACCTCTCTTTGGATAGAGAAAGAGTAGGCTGTATGTTTGGAGCAA	420
Qy	421	GGATCAATATTTTTCATTGACCTGGTTAATATCAAGAAATTTCCAAAAGATTGTGTGCC	480
Db	421	GGATCAATATTTTTCATTGACCTGGTTAATATCAAGAAATTTTCCAAAAGATTGTGTGCC	480
Qy	481	AGTATCTTACCAAGAAAGATGAATGCAAGTGGGCTGGAAAAGATCTCGAAAAGATG	540
Db	481	AGTATCTTACCAAGAAAGATGAATGCAAGTGGGCTGGAAAAGATCTCGAAAAGATG	540

Db	461	AGTATCTTACACGAGAAAGATGATGCAATGGGCTGAAAAAGCATCTGAAAGAAATG	540
Qy	541	TGCTAATTTCAATCAAGGACTTTAAGGCAATATATACAGCTCACTTGTACGCTGTGAAAC	600
Db	541	TGCTAATTTCAATCAAGGACTTTAAGGCAATATATACAGCTCACTTGTACGCTGTGAAAC	600
Qy	601	GGGGGCTTTTCAATCCAAATTTGCAACCTAATTTGAAATTGGAACATCATCTCTGGAGCAATAT	660
Db	601	GGGGGCTTTTCAATCCAAATTTGCAACCTAATTTGAAATTGGAACATCATCTCTGGAGCAATAT	660
Qy	661	TTTTAGCTGAGAACTCACATTTTGGAAAAAGCGCGGTGGAAAGAGTCCATATGACCTTAA	720
Db	661	TTTTAGCTGAGAACTCACATTTTGGAAAAAGCGCGGTGGAAAGAGTCCATATGACCTTAA	720
Qy	721	GCTGCTGACAGCATCCCTTTTAAATAGATGAGAAATTATATCTCTGAACTGACGTGATTT	780
Db	721	GCTGCTGACAGCATCCCTTTTAAATAGATGAGAAATTATATCTCTGAACTGACGTGATTT	780
Qy	781	TATGGGGGGAACATTTGCTATCTTCCGAACCTCTTGGGACCAACCACTCCAAATCAGACAG	840
Db	781	TATGGGGGGAACATTTGCTATCTTCCGAACCTCTTGGGACCAACCACTCCAAATCAGACAG	840
Qy	841	GAAGCATGATTCACAGGTGGCTCAATGATCCAAAGTTCTATATGTGGCCACCTCATCTCAGA	900
Db	841	GAAGCATGATTCACAGGTGGCTCAATGATCCAAAGTTCTATATGTGGCCACCTCATCTCAGA	900
Qy	901	GAGTGACAAATCCTGAAAGATGACAAAGTATCTTTTCTTCCGTGAAAAATGCATATGATGG	960
Db	901	GAGTGACAAATCCTGAAAGATGACAAAGTATCTTTTCTTCCGTGAAAAATGCATATGATGG	960
Qy	961	AGAACACTCTTGAAAAAGCTACTCACACGTAGAAATAGTGTAGATATTCAGAAATGACTTTGG	1020
Db	961	AGAACACTCTTGAAAAAGCTACTCACACGTAGAAATAGTGTAGATATTCAGAAATGACTTTGG	1020
Qy	1021	AGGGCACAAGAGTCTGGTGAATTAATGAGCAACAATTCCTCAAGCGTGGTCTGATTTGCTC	1080
Db	1021	AGGGCACAAGAGTCTGGTGAATTAATGAGCAACAATTCCTCAAGCGTGGTCTGATTTGCTC	1080
Qy	1081	AGTGCAGAGTCCAAATGGCATTGACACTCATTTTGTATGATACCTGCAGAGATGATTCCTAAT	1140
Db	1081	AGTGCAGAGTCCAAATGGCATTGACACTCATTTTGTATGATACCTGCAGAGATGATTCCTAAT	1140
Qy	1141	GAACCTTAAATCTCTAAAAATCCAGTTGTATATGAGTGTTTTACAGCTTCCAGTAATAT	1200
Db	1141	GAACCTTAAATCTCTAAAAATCCAGTTGTATATGAGTGTTTTACAGCTTCCAGTAATAT	1200
Qy	1201	TTTCAAGGGATCAGCCGTGTATGTATATGATGATGATGATGATGATGATGATGATGATGATG	1260
Db	1201	TTTCAAGGGATCAGCCGTGTATGTATATGATGATGATGATGATGATGATGATGATGATGATG	1260
Qy	1261	TCCATATGCCCAAGGATGAGACCCAACTATCAATGGGTGCTTATCAGGAAGAGTCC	1320
Db	1261	TCCATATGCCCAAGGATGAGACCCAACTATCAATGGGTGCTTATCAGGAAGAGTCC	1320
Qy	1321	CTATTCACGGCCAGGAATTTGTCCACAGAAAACATTTGGTGGTTTGTACTTCAAAAGGA	1380
Db	1321	CTATTCACGGCCAGGAATTTGTCCACAGAAAACATTTGGTGGTTTGTACTTCAAAAGGA	1380
Qy	1381	CCTTCCTGATATGTATATACCTTTGGCAAGAAAGCATCAGCCATGTATCAATCCAGTGT	1440
Db	1381	CCTTCCTGATATGTATATACCTTTGGCAAGAAAGCATCAGCCATGTATCAATCCAGTGT	1440
Qy	1441	TCCTATGAAACATGGCCCAATAGATCAAAAGGATATTAATTAATTAATTAATTAATTAAT	1500
Db	1441	TCCTATGAAACATGGCCCAATAGATCAAAAGGATATTAATTAATTAATTAATTAATTAAT	1500
Qy	1501	TGTGTGATGACCGAGTGTGATGCAAGAAAGTATGACAGTATGATGTATATATATATATAT	1560
Db	1501	TGTGTGATGACCGAGTGTGATGCAAGAAAGTATGACAGTATGATGTATATATATATATAT	1560
Qy	1561	TGTTGGGACGGTTCTTAAAGTATTAATTCCTAAGAGACTTGTATGATTTAGAAAGA	1620
Db	1561	TGTTGGGACGGTTCTTAAAGTATTAATTCCTAAGAGACTTGTATGATTTAGAAAGA	1620

```

Db      1561 TGTGGAGCCGTTCTTAAAGTAGTTCAATTCCTTAAGAGAGACTGGATGATTAGAAAGA 1620
Qy      1621 GGTTCGTGGAAGAAATGACAGTTTTTCGGGAGCCGACTGCTATTTGACAAATGAGACT 1680
Db      1621 GGTTCGTGGAAGAAATGACAGTTTTTCGGGAGCCGACTGCTATTTGACAAATGAGACT 1680
Qy      1681 TTCCATAGAGCAGCAACAATATATATTTGTTCAACGGCTGGGGTTGCCAGCTCCCTTT 1740
Db      1681 TTCCATAGAGCAGCAACAATATATATTTGTTCAACGGCTGGGGTTGCCAGCTCCCTTT 1740
Qy      1741 ACACCGGTGTATATTTACGGGAAAGCGTGTCTGATGTGTCTGCTGCCCGGAGACCTTTA 1800
Db      1741 ACACCGGTGTATATTTACGGGAAAGCGTGTCTGATGTGTCTGCTGCCCGGAGACCTTTA 1800
Qy      1801 CTGTCTTGGGATGTTCTGCAATGTTCTCGATTTTCCACTGCAAGAGACGACAAAG 1860
Db      1801 CTGTCTTGGGATGTTCTGCAATGTTCTCGATTTTCCACTGCAAGAGACGACAAAG 1860
Qy      1861 ACGACAGATATTAAGAAATGAGAGCCCACTGACTGTTTCAAGCTTACACATGATTA 1920
Db      1861 ACGACAGATATTAAGAAATGAGAGCCCACTGACTGTTTCAAGCTTACACATGATTA 1920
Qy      1921 TCACATGCGCCACAGCCCTGAAAGAGAGATCATCTATGTGTAGAGAAATGATGACACTT 1980
Db      1921 TCACATGCGCCACAGCCCTGAAAGAGAGATCATCTATGTGTAGAGAAATGATGACACTT 1980
Qy      1981 TTAGGATGACAGTCCGAGAGTGCAGAGAGCGCTGTGTTTGGCAATTCAGAGGCGAA 2040
Db      1981 TTAGGATGACAGTCCGAGAGTGCAGAGAGCGCTGTGTTTGGCAATTCAGAGGCGAA 2040
Qy      2041 TGAAGAGCGAAAGAGAGATCAGATGATGATCATATCATCAGAGACGATCAAGGCT 2100
Db      2041 TGAAGAGCGAAAGAGAGATCAGATGATGATCATATCATCAGAGACGATCAAGGCT 2100
Qy      2101 TCTGTAGTGTGTCTACACAGAGAGATTCAGGCAATTAACCTCTGCCATGCGGTGGAACA 2160
Db      2101 TCTGTAGTGTGTCTACACAGAGAGATTCAGGCAATTAACCTCTGCCATGCGGTGGAACA 2160
Qy      2161 TGGGTTCTATACAACTTCTTAAGTAACTCCGAGAGTCAATGACAGAGCAATTTGGA 2220
Db      2161 TGGGTTCTATACAACTTCTTAAGTAACTCCGAGAGTCAATGACAGAGCAATTTGGA 2220
Qy      2221 AGAAGCTTCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2280
Db      2221 AGAAGCTTCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2280
Qy      2281 CATGACACTTACCGAGAGGTCTGTGTACAGAGACTTTCATGACGCTCATCAACACCCAA 2340
Db      2281 CATGACACTTACCGAGAGGTCTGTGTACAGAGACTTTCATGACGCTCATCAACACCCAA 2340
Qy      2341 TCTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
Db      2341 TCTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
Qy      2401 GCAAGAGCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2460
Db      2401 GCAAGAGCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2460
Qy      2461 GAAAGGTGAAACAGAGAGAGCCAGAAATTTGAGAGGAGACCCAGAGGTGTCTGAGTGC 2520
Db      2461 GAAAGGTGAAACAGAGAGAGCCAGAAATTTGAGAGGAGACCCAGAGGTGTCTGAGTGC 2520
Qy      2521 ATTACCTTAGAACTCTCAAACTTCCCTAGACAAATTAACCTGAGAAAAACAA 2580
Db      2521 ATTACCTTAGAACTCTCAAACTTCCCTAGACAAATTAACCTGAGAAAAACAA 2580
Qy      2581 TGCATATATCAATGAACTTTTTCATGCAATTAATGATGATGATGATGATGATGATGATGATG 2640
Db      2581 TGCATATATCAATGAACTTTTTCATGCAATTAATGATGATGATGATGATGATGATGATGATG 2640
Qy      2641 AGCTAGTTTCCACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2700
Db      2641 AGCTAGTTTCCACCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2700

```

```

Qy      2701 CCTATACC 2709
Db      2701 CCTATACC 2709

RESULT 2
US-10-097-340-283
; Sequence 283, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNANAVARAPU
; APPLICANT: Sebastian HOERSCHE
; APPLICANT: Shubhangi KAMATKAR
; APPLICANT: Steve G. KOVARS
; APPLICANT: Rachel E. MEYERS
; APPLICANT: Michael MORRISSEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIRBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANDT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: MRI-030
; CURRENT FILING DATE: 2002-03-14
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 283
; LENGTH: 2530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-097-340-283

Query Match      92.6%; Score 2508; DB 14; Length 2530;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY	433	TTGATTTCCACCTTGCTTAATATCAAGATTTTCAAAAAGTTGTGTGGCCAGTATCTTACAC	492
Db	249	TTTCATTCCGCTGGTAAATATCAAGATTTTCAAAAAGTTGTGTGGCCAGTATCTTACAC	308
QY	493	CAGAAGAGTATGATCAAGTGGGCTGGAAMAACATCTCGAAGAATGTGCTAATTTTCAT	552
Db	309	CAGAAGAGTATGATCAAGTGGGCTGGAAMAACATCTCGAAGAATGTGCTAATTTTCAT	368
QY	553	CAGGTACTTAAAGCATATATCAACTTGTACCTGTGGACGGGGGCTTTTCA	612
Db	369	CAGGTACTTAAAGCATATATCAACTTGTACCTGTGGACGGGGGCTTTTCA	428
QY	613	TCGATTTGGACCTTACATGAAATTTGACATCATCTGAGGACAAATTTTTTAAGCTGGA	672
Db	429	TCGATTTGGACCTTACATGAAATTTGACATCATCTGAGGACAAATTTTTTAAGCTGGA	488
QY	673	GAACTCACATTTTGAAMAAGCGCGGTGGAGAAGTCATATGACCTTAAGCTGTGACAGC	732
Db	489	GAACTCACATTTTGAAMAAGCGCGGTGGAGAAGTCATATGACCTTAAGCTGTGACAGC	548
QY	733	ATCCCTTTTAAATAGATGAGAAATTATCTGTGGAACTGACGTGATTTTATGGGGCCAGA	792
Db	549	ATCCCTTTTAAATAGATGAGAAATTATCTGTGGAACTGACGTGATTTTATGGGGCCAGA	608
QY	793	CTTTGCTATCTTCCCAACTCTTGGGACACCAACCCTAATCAGGACAGACAGATATTC	852
Db	609	CTTTGCTATCTTCCCAACTCTTGGGACACCAACCCTAATCAGGACAGACAGATATTC	668
QY	853	CAGGTGGCTCATGATCCAAAGTTCAATATGTGCCCACTCATCTGAGAAGTGAATATCC	912
Db	669	CAGGTGGCTCATGATCCAAAGTTCAATATGTGCCCACTCATCTGAGAAGTGAATATCC	728
QY	913	TGAAGATGACAAAGTATCTTTTCTTCGCTGAATAATCAATAGATGAGAAACCTCTGG	972
Db	729	TGAAGATGACAAAGTATCTTTTCTTCGCTGAATAATCAATAGATGAGAAACCTCTGG	788
QY	973	AAAGCTTCTACGCTAGAAATAGTCAATATGCAAGAATGACTTTGGAGGGCACAGAA	1032
Db	789	AAAGCTTCTACGCTAGAAATAGTCAATATGCAAGAATGACTTTGGAGGGCACAGAA	848
QY	1033	TCTGGTGAATTAATGAGACAACTTCTCAAAAGCTGTGATTTGGTCTAGTGCCAGGTCC	1092
Db	849	TCTGGTGAATTAATGAGACAACTTCTCAAAAGCTGTGATTTGGTCTAGTGCCAGGTCC	908
QY	1093	AAATGGCATTTGACACTCAATTTTGTGATGAACTGCAGAGTATTTCTTAATGAACCTTAA	1152
Db	909	AAATGGCATTTGACACTCAATTTTGTGATGAACTGCAGAGTATTTCTTAATGAACCTTAA	968
QY	1153	TCCTTAAATCCAGTTGTATATGAGATGTTTACGACTTCCAGTAAACATTTTCAAGGATC	1212
Db	969	TCCTTAAATCCAGTTGTATATGAGATGTTTACGACTTCCAGTAAACATTTTCAAGGATC	1028
QY	1213	AGCGGTGTATGTATAGCATGAGTGTGAGAGGGTGTCTTCTGTCTCATATGCCA	1272
Db	1029	AGCGGTGTATGTATAGCATGAGTGTGAGAGGGTGTCTTCTGTCTCATATGCCA	1088
QY	1273	CAGGATGACCCAACTATCAATGGGTCTTATCAAGAAAGTCCCTTATCCAGGCTC	1332
Db	1089	CAGGATGACCCAACTATCAATGGGTCTTATCAAGAAAGTCCCTTATCCAGGCTC	1148
QY	1333	AGGAATCTGTCCAGCAAAACATTTGTGTGTTTGAACCTTCAAGAGACCTTCTGATGA	1392
Db	1149	AGGAATCTGTCCAGCAAAACATTTGTGTGTTTGAACCTTCAAGAGACCTTCTGATGA	1208
QY	1393	TGTTATTAACCTTTGCAAGAAGTCATCCAGCATGTACATCCAGTGTTTCTTATGAACA	1452
Db	1209	TGTTATTAACCTTTGCAAGAAGTCATCCAGCATGTACATCCAGTGTTTCTTATGAACA	1268
QY	1453	TCGCCCAATAGTATCAAAAAGGATGTAAATTTATCAATTTACAAAATTTGTCTGATGACG	1512
Db	1269	TCGCCCAATAGTATCAAAAAGGATGTAAATTTATCAATTTACAAAATTTGTCTGATGACG	1328

QY	1513	AGTGGATGCAGAAAGATGCAAGATGATGATGTTATTCGGAACAGATGTTGGGACCGT	1517
Db	1339	AGTGATGCAGAAAGATGCAAGATGATGATGTTATTCGGAACAGATGTTGGGACCGT	1388
QY	1573	TCTTAAATAGTTTCATTTCTTAAGAGACCTTGGATATTTAGAGAGTTCCTGCTGGA	1632
Db	1389	TCTTAAATAGTTTCATTTCTTAAGAGACCTTGGATATTTAGAGAGTTCCTGCTGGA	1448
QY	1633	AGAAATGACAGTTTTTCGGGAACCGACTGCTATTTTCAGCAATGAGAGCTTTCOACTAAGCA	1692
Db	1449	AGAAATGACAGTTTTTCGGGAACCGACTGCTATTTTCAGCAATGAGAGCTTTCOACTAAGCA	1508
QY	1653	GCAACAACTATATATATGTTGTTCAACGGCTGGGGTTGCCAGCTCCCTTTACACCGTGGA	1755
Db	1509	GCAACAACTATATATATGTTGTTCAACGGCTGGGGTTGCCAGCTCCCTTTACACCGTGGA	1568
QY	1753	TATTTACGGGAAAGCGTGTGCTGATGTTGCTCGCCGAGAACCTTACTGCTGGTGGGA	1812
Db	1569	TATTTACGGGAAAGCGTGTGCTGATGTTGCTCGCCGAGAACCTTACTGCTGGTGGGA	1628
QY	1813	TGGTTCTGCAATGTTCTCGCTATTTTTCCCACTGCAGAGAGCGCACAGACGACAAGATAT	1877
Db	1629	TGGTTCTGCAATGTTCTCGCTATTTTTCCCACTGCAGAGAGCGCACAGACGACAAGATAT	1688
QY	1873	AAGAAATGGAACCCACTGACTCTACTGTTCAGACTTACACATGATATATCACACTGGCCA	1932
Db	1689	AAGAAATGGAACCCACTGACTCTACTGTTCAGACTTACACATGATATATCACACTGGCCA	1748
QY	1933	CAGCCCTGAAGAGAAATCATCTATGAGGTGAGAAATGTAGCAATTTTGGATGATGAG	1992
Db	1749	CAGCCCTGAAGAGAAATCATCTATGAGGTGAGAAATGTAGCAATTTTGGATGATGAG	1808
QY	1993	TCCGAGTCGAGAGAGAGCGCTGCTCTATTTGGCAATTCAGAGGCGAAATGAAGACGAA	2052
Db	1809	TCCGAGTCGAGAGAGAGCGCTGCTCTATTTGGCAATTCAGAGGCGAAATGAAGACGAA	1868
QY	2053	AGAAAGATCATGATGATGATATCATATATCATAAGACAGATTAAGGCTTCTGCTACGTAG	2112
Db	1869	AGAAAGATCATGATGATGATATCATATATCATAAGACAGATTAAGGCTTCTGCTACGTAG	1928
QY	2113	TCTTCAACAGAGAAATTCAGGCAATTCAGTCAATTCCTGTCAGTCCGCTGGACATGGGTTCAATCA	2172
Db	1929	TCTTCAACAGAGAAATTCAGGCAATTCAGTCAATTCCTGTCAGTCCGCTGGACATGGGTTCAATCA	1988
QY	2173	AACTCTTCTTAAGGTAAACCCCTGGAAGTCAATTGACACAGAGCATTTGGAGAACTTCTTCA	2232
Db	1989	AACTCTTCTTAAGGTAAACCCCTGGAAGTCAATTGACACAGAGCATTTGGAGAACTTCTTCA	2048
QY	2233	TAAAGATGATATGAGATGAGTGCCTTAAGACCAAGAAATGCTCAATGATGACACTTGA	2292
Db	2049	TAAAGATGATATGAGATGAGTGCCTTAAGACCAAGAAATGCTCAATGATGACACTTGA	2108
QY	2293	CCAGAAGGTCTGTGATCAGAGACTTCATGACAGCTCATCAACCAACCCCAATCTCAACAGAT	2352
Db	2109	CCAGAAGGTCTGTGATCAGAGACTTCATGACAGCTCATCAACCAACCCCAATCTCAACAGAT	2168
QY	2353	GGATGAGTCTGTGTAACAAGTTTGGAAAAAGGACCGAAAAACAAGTCGGCAAGGCGCAGG	2412
Db	2169	GGATGAGTCTGTGTAACAAGTTTGGAAAAAGGACCGAAAAACAAGTCGGCAAGGCGCAGG	2228
QY	2413	ACATATCCCGAGGGAACAGTAAACAATGGAAGCACTTACAGAAATATGAAGAGGTAGAAA	2472
Db	2229	ACATATCCCGAGGGAACAGTAAACAATGGAAGCACTTACAGAAATATGAAGAGGTAGAAA	2288
QY	2473	CAGAGAGACCCACGAATTTTGAAGGGGACCCACGAGAGTGTCTGAGCTGATTAACCTCTAGA	2532
Db	2289	CAGAGAGACCCACGAATTTTGAAGGGGACCCACGAGAGTGTCTGAGCTGATTAACCTCTAGA	2348
QY	2533	AACTCAAAACAATGAGAAACTTGGCTTGAACAATATCTGAAAAACAATGCAATATACAT	2592
Db	2349	AACTCAAAACAATGAGAAACTTGGCTTGAACAATATCTGAAAAACAATGCAATATACAT	2408
QY	2593	GAACTTTTTCATGCGCATATATGAGATGTTTACAGATGAGGGAATTCAGCTGAGTTCCA	2652


```

Db      1689 AAGAAATGAGAGACCCCTGACTGCTTCAAGCTTACACCAATGATATACCATGACCATGAGCA 1748
Qy      1933 CAGCCCTGAAGAGAGATATCTATGCTGAGAGATAGATAGACATTTTGGAAAGCAG 1992
Db      1749 CAGCCCTGAAGAGAGATATCTATGCTGAGAGATAGATAGACATTTTGGAAAGCAG 1808
Qy      1993 TCCGAAGTGCAGAGAGAGCGCTGCTTATTTGGCAATTCAGAGGCGAAATGAAAGCGAAA 2052
Db      1809 TCCGAAGTGCAGAGAGAGCGCTGCTTATTTGGCAATTCAGAGGCGAAATGAAAGCGAAA 1868
Qy      2053 AGAAGAGATCAGATGATGATGATATATCATCAGAGCAATGCAAGGCTTCTGCTACGTAG 2112
Db      1869 AGAAGAGATCAGATGATGATGATATATCATCAGAGCAATGCAAGGCTTCTGCTACGTAG 1928
Qy      2113 TCTACAAAGAGAGATTCAGAGGCAATTCAGGCTGCGAGGCTGAGGCTGAGGCTTCTATATA 2172
Db      1929 TCTACAAAGAGAGATTCAGAGGCAATTCAGGCTGCGAGGCTGAGGCTGAGGCTTCTATATA 1988
Qy      2173 AACTCTTCTTAAGTAAACCTGGAATGATGACAGAGCAATTTGGAAGAACTTCTTCA 2232
Db      1989 AACTCTTCTTAAGTAAACCTGGAATGATGACAGAGCAATTTGGAAGAACTTCTTCA 2048
Qy      2233 TAAAGATGATGATGATGATGCTCTTAAGACCAAGAAATGCTCAATAGCATGACACTAG 2292
Db      2049 TAAAGATGATGATGATGATGCTCTTAAGACCAAGAAATGCTCAATAGCATGACACTAG 2108
Qy      2293 CGAAGAGTCTGCTGATAGAGACTTCAAGGCTGATGACCAAGGCTGAGGCTTCTTCAAGCAT 2352
Db      2109 CGAAGAGTCTGCTGATAGAGACTTCAAGGCTGATGACCAAGGCTGAGGCTTCTTCAAGCAT 2168
Qy      2353 GGATGAGTCTGCTGATAGAGACTTCAAGGCTGATGACCAAGGCTGAGGCTTCTTCAAGCAT 2412
Db      2169 GGATGAGTCTGCTGATAGAGACTTCAAGGCTGATGACCAAGGCTGAGGCTTCTTCAAGCAT 2228
Qy      2413 ACATATCCCAAGGAGACAGTAACTGAAAGCACTTCAAGAAATTAAGAAAGTAAAGAA 2472
Db      2229 ACATATCCCAAGGAGACAGTAACTGAAAGCACTTCAAGAAATTAAGAAAGTAAAGAA 2288
Qy      2473 CAGGAGGAGCCAGATTTGAGAGGCGACCCAGAGGCTGCTGAGCTTCAAGCTTCTTCA 2532
Db      2289 CAGGAGGAGCCAGATTTGAGAGGCGACCCAGAGGCTGCTGAGCTTCAAGCTTCTTCA 2348
Qy      2533 AACCTCAAAAGATAGAACTTCTGCTAGACCAATTAAGTGAAGAAACCAATGCAATATACAT 2592
Db      2349 AACCTCAAAAGATAGAACTTCTGCTAGACCAATTAAGTGAAGAAACCAATGCAATATACAT 2408
Qy      2593 GAACCTTTTTCATGCGATATGATGATGATTTTACATGCTGGAATTCAGCTGAGTTCCA 2652
Db      2409 GAACCTTTTTCATGCGATATGATGATGATTTTACATGCTGGAATTCAGCTGAGTTCCA 2468
Qy      2653 CCAATTAATTAATTAATTCATGATGATGATTTTCTTCTTATAGGCTTTTCTT 2700
Db      2469 CCAATTAATTAATTAATTCATGATGATGATTTTCTTCTTATAGGCTTTTCTT 2516

RESULT 4
US-09-864-761-30992
; Sequence 30992, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME- DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aemica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456

```

```

; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30992
; LENGTH: 456
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006322.2
; OTHER INFORMATION: EXPRESSED IN LONG, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.55
; OTHER INFORMATION: EST HUMAN HIT: BF700780.1, EVALU0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALU0.00e+00
; OTHER INFORMATION: NT HIT: g111421514, EVALU0.00e+00
US-09-864-761-30992

Query Match      15.0%; Score 405; DB 9; Length 456;
Best Local Similarity 99.8%; Pred.No. 1.2e-211;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy      2060 ATCAGAGTGTGATGATCATATATCATCAGAGACAGATCAAGGCTTCTGCTACGATGCTACAA 2119
Db      1 ATCAGAGTGTGATGATCATATATCATCAGAGACAGATCAAGGCTTCTGCTACGATGCTACAA 60
Qy      2120 CAGAAGGATTCAGGCAATTAACCTGCGCATCGGCTGAGAAATGGGTTCTATCAAACTCTT 2179
Db      61 CAGAAGGATTCAGGCAATTAACCTGCGCATCGGCTGAGAAATGGGTTCTATCAAACTCTT 120
Qy      2180 CTTAAGGTAACCTGGAAGTCAATGACACAGAGCAATTTGGAAGAACTTTCTATTAAGAT 2239
Db      121 CTTAAGGTAACCTGGAAGTCAATGACACAGAGCAATTTGGAAGAACTTTCTATTAAGAT 180
Qy      2240 GATGATGAGATGGCTCTTAAGACCAAGAAATGCTCAATAGCATGACACCTTACGCAAG 2299
Db      181 GATGATGAGATGGCTCTTAAGACCAAGAAATGCTCAATAGCATGACACCTTACGCAAG 240
Qy      2300 GTCTGGTACAGAGCTTCATGACAGCAAGCAACCAATCTCAACAGATGATGAG 2359
Db      241 GTCTGGTACAGAGCTTCATGACAGCAAGCAACCAATCTCAACAGATGATGAG 300
Qy      2360 TTCTGTGAACAAGTTTGGAAAAAGGAGCGAAGAAACAGCTGCGCAAGGCCAGACATACC 2419

```

Db 301 TTCTGTGAACAAGTTGAGAAAGGAGCCGCAAGGCGCAGACATACC 360
Qy 2420 CCAGGACACGATTAACAATGAGACCTTACAGAAATTAAGAAAGTGAAGACGAGG 2479
Db 361 CCGGGAAACGTAACAATGAGACCTTACAGAAATTAAGAAAGTGAAGACGAGG 420
Qy 2480 ACCCAGAAATTTGAGAGGACACCCAGAGTGTCTGA 2515
Db 421 ACCCAGAAATTTGAGAGGACACCCAGAGTGTCTGA 456

RESULT 5

US-09-922-217-736
; Sequence 736, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922,217
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq For Windows Version 4.0
; SEQ ID NO 736
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 57
; OTHER INFORMATION: n = A,T,C or G
US-09-922-217-736

Query Match 11.2%; Score 303; DB 9; Length 354;
Best Local Similarity 99.7%; Pred. No. 1.5e-155;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2098 CCTCTGCTAGTGTCTTACAAAGAGATTACAGCAATTAACCTGCGCATGGGGTGA 2157
Db 1 CCTCTGCTAGTGTCTTACAAAGAGATTACAGCAATTAACCTGCGCATGGGGTGA 60
Qy 2158 ACATGGGTTTCAATCAAACTCTTCTTAAGTTAACTGGAAGTCATTGACAGAGCATTT 2217
Db 61 ACATGGGTTTCAATCAAACTCTTCTTAAGTTAACTGGAAGTCATTGACAGAGCATTT 120
Qy 2218 GGAAGAACTTCTTCAATAAAGATGATGAGAGTGGCTCTTAAGACCAAGAAATGTCAA 2277
Db 121 GGAAGAACTTCTTCAATAAAGATGATGAGAGTGGCTCTTAAGACCAAGAAATGTCAA 180
Qy 2278 TAGCATGACCTTACCGCAGAGGTGTGTACAGAACTTACGAGCTCATCAACCAACC 2337
Db 181 TAGCATGACCTTACCGCAGAGGTGTGTACAGAACTTACGAGCTCATCAACCAACC 240
Qy 2338 CAATCTCAACAGATGATGAGTGTGTGAACAAGTTGAAAAAGGACCGAAAAACAAG 2397
Db 241 CAATCTCAACAGATGATGAGTGTGTGAACAAGTTGAAAAAGGACCGAAAAACAAG 300
Qy 2398 TCGGCAAAAGCCAGAGCATACCCCGAGGAAAGTAAACATGAAGACCTTACA 2451
Db 301 TCGGCAAAAGCCAGAGCATACCCCGAGGAAAGTAAACATGAAGACCTTACA 354

RESULT 6

US-09-833-263-736
; Sequence 736, Application US/09833263
; Patent No. US20020110547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aijun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833,263
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq For Windows Version 3.0
; SEQ ID NO 736
; LENGTH: 354
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(354)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-263-736

Query Match 11.2%; Score 303; DB 10; Length 354;
Best Local Similarity 99.7%; Pred. No. 1.5e-155;
Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2098 CCTCTGCTAGTGTCTTACAAAGAGATTACAGCAATTAACCTGCGCATGGGGTGA 2157
Db 1 CCTCTGCTAGTGTCTTACAAAGAGATTACAGCAATTAACCTGCGCATGGGGTGA 60
Qy 2158 ACATGGGTTTCAATCAAACTCTTCTTAAGTTAACTGGAAGTCATTGACAGAGCATTT 2217
Db 61 ACATGGGTTTCAATCAAACTCTTCTTAAGTTAACTGGAAGTCATTGACAGAGCATTT 120
Qy 2218 GGAAGAACTTCTTCAATAAAGATGATGAGAGTGGCTCTTAAGACCAAGAAATGTCAA 2277
Db 121 GGAAGAACTTCTTCAATAAAGATGATGAGAGTGGCTCTTAAGACCAAGAAATGTCAA 180
Qy 2278 TAGCATGACCTTACCGCAGAGGTGTGTACAGAACTTACGAGCTCATCAACCAACC 2337
Db 181 TAGCATGACCTTACCGCAGAGGTGTGTACAGAACTTACGAGCTCATCAACCAACC 240
Qy 2338 CAATCTCAACAGATGATGAGTGTGTGAACAAGTTGAAAAAGGACCGAAAAACAAG 2397
Db 241 CAATCTCAACAGATGATGAGTGTGTGAACAAGTTGAAAAAGGACCGAAAAACAAG 300
Qy 2398 TCGGCAAAAGCCAGAGCATACCCCGAGGAAAGTAAACATGAAGACCTTACA 2451
Db 301 TCGGCAAAAGCCAGAGCATACCCCGAGGAAAGTAAACATGAAGACCTTACA 354

RESULT 7

US-10-025-380-736
; Sequence 736, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aijun


```

: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzei, David K

```


APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 27943
LENGTH: 172
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
OTHER INFORMATION: NT HIT: g11421514, EVALUE 8.00e-89
OTHER INFORMATION: EST HUMAN HIT: BE694785.1, EVALUE 2.30e+00
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
US-09-864-761-27943
Query Match 5.9%; Score 160; DB 9; Length 172;
Best Local Similarity 100.0%; Pred. No. 6,7e-77;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cy 310 AGAATGTTGAATCAACAAATGATCATCTTCAATGCTTGGCCAAACAGCTCCAGTTA 369
Db 160 AGAATGTTGAATCAACAAATGATCATCTTCAATGCTTGGCCAAACAGCTCCAGTTA 101
Cy 370 TCATACCTCTCTTTGGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 429
Db 100 TCATACCTCTCTTTGGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCAT 41

Cy 430 ATTTTCATTCGACCTGGTATATATCAAGATTTTCAAG 469
Db 40 ATTTTCATTCGACCTGGTATATATCAAGATTTTCAAG 1
RESULT 10
US-09-864-761-11363/c
Sequence 11363, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 11363
LENGTH: 484
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
US-09-864-761-11363
Query Match 5.9%; Score 160; DB 9; Length 484;
Best Local Similarity 100.0%; Pred. No. 7.4e-77;


```
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 28381
;; LENGTH: 123
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC006322.2
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.46
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.54
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56
;; OTHER INFORMATION: EST HUMAN HIT: BF667677.1, EVALUE 3.00e-63
;; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 1.00e-18
;; OTHER INFORMATION: NT HIT: g111421514, EVALUE 2.00e-63
US-09-864-761-28381

Query Match          4.5%; Score 123; DB 9; Length 123;
Best Local Similarity 100.0%; Pred. No. 1.5e-56;
Matches 123; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1937 CCTGAGAGAGATCATCTATGCTGTAGAGATAGTACACATTTTGGAAATGCAGTCCG 1996
Db 1 CCTGAGAGAGATCATCTATGCTGTAGAGATAGTACACATTTTGGAAATGCAGTCCG 60

Cy 1997 AAGTCGACAGAGCGGCTGCTATTGGCAATTCCAGAGCGCAATGAAGCGAAAGAA 2056
Db 61 AAGTCGACAGAGCGGCTGCTATTGGCAATTCCAGAGCGCAATGAAGCGAAAGAA 120

Cy 2057 GAG 2059
Db 121 GAG 123

RESULT 13
US-09-864-761-6056/c
;; Sequence 6056, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: Aecomica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; PRIOR FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263,6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
```

```
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 6056
;; LENGTH: 456
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC004848.1
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
US-09-864-761-6056

Query Match          4.5%; Score 122; DB 9; Length 456;
Best Local Similarity 100.0%; Pred. No. 5.9e-56;
Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 532 GAAGAATGTGCTAATTCATCAAGTACTTAAGCATATATAGACTCACTTGTACGC 591
Db 407 GAAGAATGTGCTAATTCATCAAGTACTTAAGCATATATAGACTCACTTGTACGC 348

Cy 592 CTGTGAAAGCGGCGCTTTCATCCAAATTGGCACTTCACTTGAATTGACATCACTCTGA 651
Db 347 CTGTGAAAGCGGCGCTTTCATCCAAATTGGCACTTGAATTGACATCACTCTGA 288

Cy 652 GG 653
Db 287 GG 286

RESULT 14
US-09-864-761-22893/c
;; Sequence 22893, Application US/09864761
;; Patent No. US20020048763A1
;; GENERAL INFORMATION:
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: Aecomica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; PRIOR FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263,6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
```

PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
SEQ ID NO 22893
LENGTH: 101
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO ACC04848.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
OTHER INFORMATION: NT HIT: g111421514, EVALUE 3.00e-50
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 4.00e-15
US-09-864-761-22893

Query Match 3.7%; Score 101; DB 9; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.8e-44;
Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 533 AAGATGTCATTAATTCATCAAGTACTTAAGGCATATATACAGACTGTTACGCC 592
DB 101 AAGATGTCATTAATTCATCAAGTACTTAAGGCATATATACAGACTGTTACGCC 42
QY 593 TGTGACGGGGGCTTTTCATCCAAATTTGCACCTACATTGA 633
DB 41 TGTGACGGGGGCTTTTCATCCAAATTTGCACCTACATTGA 1

RESULT 15
US-09-774-490-6
Sequence 6, Application US/09774490
Patent No. US20010034332A1
GENERAL INFORMATION:
APPLICANT: Jin, Shengfang
TITLE OF INVENTION: RESISTANCE SEQUENCES AND USES THEREOF
FILE REFERENCE: 07334-138001
CURRENT FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: US 60/179,191
PRIOR FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 2137
TYPE: DNA
ORGANISM: Homo sapiens
US-09-774-490-6

Query Match 2.6%; Score 71; DB 9; Length 2137;

Best Local Similarity 100.0%; Pred. No. 7.8e-28;
Matches 71; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTATTTATGATGTTAACAAGCTAGTAATGATGACCGACGCGGGGTGCG 60
DB 17 AATCTTATTTATGATGTTAACAAGCTAGTAATGATGACCGACGCGGGGTGCG 76
QY 61 ACCCAGCGCTC 71
DB 77 ACCCAGCGCTC 87

RESULT 16
US-10-027-632-238542
Sequence 238542, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US/10/027,632
PRIOR FILING DATE: 2000-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 238542
LENGTH: 520
TYPE: DNA
ORGANISM: Human
US-10-027-632-238542

Query Match 2.5%; Score 67; DB 13; Length 520;
Best Local Similarity 100.0%; Pred. No. 1.1e-25;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1850 AGACGACAGACGACAGATATATAGAAATGAGAGACCACTGACTCTGTTACACTTA 1909
DB 364 AGACGACAGACGACAGATATATAGAAATGAGAGACCACTGACTCTGTTACACTTA 423
QY 1910 CACCATG 1916
DB 424 CACCATG 430

RESULT 17
US-10-027-632-238543
Sequence 238543, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US/10/027,632
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 238543
LENGTH: 520
TYPE: DNA
ORGANISM: Human
US-10-027-632-238543

Query Match 2.5%; Score 67; DB 13; Length 520;
Best Local Similarity 100.0%; Pred. No. 1.1e-25;
Matches 67; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1850 AGACGACAGACGACAGATATAGAAATGGAGACCCCTGCTCAGTCTTCAAGCTTA 1909
DB 364 AGACGACAGACGACAGATATAGAAATGGAGACCCCTGCTCAGTCTTCAAGCTTA 423

QY 1910 CACCATG 1916
DB 424 CACCATG 430

RESULT 18
US-09-897-006-10/c
Sequence 10, Application US/09897006
Patent No. US20020106729A1
GENERAL INFORMATION:
APPLICANT: Bleck, Gregory
TITLE OF INVENTION: Expression Vectors
FILE REFERENCE: GALA-06415
CURRENT APPLICATION NUMBER: US/09/897,006
CURRENT FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 60/215,851
PRIOR FILING DATE: 2000-07-03
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.0
SEQ ID NO 10
LENGTH: 4661
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-09-897-006-10

Query Match 1.1%; Score 31; DB 10; Length 4661;
Best Local Similarity 100.0%; Pred. No. 8.4e-06;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATGTTAACAAGCTT 31
DB 4041 AATCTTTATTTATCGATGTTAACAAGCTT 4011

RESULT 19
US-09-897-511A-10/c
Sequence 10, Application US/09897511A
Publication No. US20030092882A1
GENERAL INFORMATION:
APPLICANT: Bremel, Robert
APPLICANT: Miller, Linda
APPLICANT: Bleck, Gregory
TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
FILE REFERENCE: GALA-06416
CURRENT APPLICATION NUMBER: US/09/897,511A
CURRENT FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 60/215,925

PRIOR FILING DATE: 2000-07-03
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.0
SEQ ID NO 10
LENGTH: 4661
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-09-897-511A-10

Query Match 1.1%; Score 31; DB 11; Length 4661;
Best Local Similarity 100.0%; Pred. No. 8.4e-06;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATGTTAACAAGCTT 31
DB 4041 AATCTTTATTTATCGATGTTAACAAGCTT 4011

RESULT 20
US-08-786-531B-5/c
Sequence 5, Application US/08786531B
Publication No. US20020015979A1
GENERAL INFORMATION:
APPLICANT: Link, Charles J.
APPLICANT: Levy, John P.
APPLICANT: Wang, Suming
APPLICANT: Seragina, Tatiana
TITLE OF INVENTION: Vehicles for Stable Transfer of Green
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: Zarley, McKee, Thomte, Voorhees & Sease
STREET: 801 Grand Suite 3200
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,531B
FILING DATE: 21-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/010371
FILING DATE: 22-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Nebel, Heidi S.
REGISTRATION NUMBER: 37,719
REFERENCE/DOCKET NUMBER: bgtri
TELECOMMUNICATION INFORMATION:
TELEPHONE: 515-288-3667
TELEFAX: 515-288-1338
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 7160 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-786-531B-5

Query Match 1.1%; Score 31; DB 8; Length 7160;
Best Local Similarity 100.0%; Pred. No. 8.8e-06;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTTATTCATGATGTAACAAGCTT 31
DB 4179 AATCTTTTATTCATGATGTAACAAGCTT 4149

RESULT 21
US-08-786-531B-6/c

Sequence 6, Application US/08786531B
Publication No. US20020015979A1
GENERAL INFORMATION:
APPLICANT: Link, Charles J.
APPLICANT: Levy, John P.
APPLICANT: Wang, Suming
APPLICANT: Seregina, Tatiana
TITLE OF INVENTION: Vehicles for Stable Transfer of Green
TITLE OF INVENTION: Fluorescent Protein Gene and Methods of Use for Same
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
STREET: 801 Grand Suite 3200
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,531B
FILING DATE: 21-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/010371
FILING DATE: 22-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Nebel, Heidi S.
REGISTRATION NUMBER: 37,719
REFERENCE/DOCKET NUMBER: hgct1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 515-288-3667
TELEFAX: 515-288-1338
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 7235 base pairs
TYPE: nucleic acid
STRADEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: other nucleic acid
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-786-531B-6

Query Match 1.1%; Score 31; DB 8; Length 7235;
Best Local Similarity 100.0%; Pred. No. 8.8e-06;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTTATTCATGATGTAACAAGCTT 31
DB 4254 AATCTTTTATTCATGATGTAACAAGCTT 4224

RESULT 22
US-09-728-422-1/c

Sequence 1, Application US/09728422
Patent No. US20020128187A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Zhou, Bing
APPLICANT: Goodrich, Kyle
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod

APPLICANT: Ren, Feiyan
APPLICANT: Zhao, Qing A.
APPLICANT: Yang, Yonchong
APPLICANT: Wehrman, Tom
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. US20020128187A1el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 787CIP2F
CURRENT APPLICATION NUMBER: US/09/728,422
CURRENT FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 10
SOFTWARE: pc_fl_genes Version 2.0
SEQ ID NO 1
LENGTH: 1313
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (547)..(1239)
US-09-728-422-1

Query Match 1.0%; Score 26; DB 10; Length 1313;
Best Local Similarity 100.0%; Pred. No. 0.0042;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 GTGTGACCCACGCGTCCGGAGTAG 80
DB 1305 GTGTGACCCACGCGTCCGGAGTAG 1280

RESULT 23
US-09-774-490-5
Sequence 5, Application US/09774490
Patent No. US20010034332A1
GENERAL INFORMATION:
APPLICANT: Jin, Shengfang
TITLE OF INVENTION: RESISTANCE SEQUENCES AND USES THEREOF
FILE REFERENCE: 07334-138001
CURRENT APPLICATION NUMBER: US/09/774,490
CURRENT FILING DATE: 2001-01-31
PRIOR APPLICATION NUMBER: US 60/179,191
PRIOR FILING DATE: 2000-01-31
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 1400
TYPE: DNA
ORGANISM: Mus musculus
US-09-774-490-5

Query Match 0.9%; Score 25; DB 9; Length 1400;
Best Local Similarity 100.0%; Pred. No. 0.015;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 TAACAAGCTTAGTATGATGCCAC 46
DB 29 TAACAAGCTTAGTATGATGCCAC 53

RESULT 24
US-10-106-698-3299
Sequence 3299, Application US/10106698
Publication No. US20030109690A1
GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
FILE REFERENCE: PA005P1
CURRENT APPLICATION NUMBER: US/10/106,698
CURRENT FILING DATE: 2002-03-27

```

; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 3239
; LENGTH: 115
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (8)..(8)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (25)..(26)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (84)..(84)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (111)..(111)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (115)..(115)
; OTHER INFORMATION: n equals a,t,g, or c
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-3299
```

```

Query Match          0.8%; Score 22; DB 14; Length 115;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 57 GTCGACCCACGCGTCGGGAGT 78
    |||||||
Db 61 GTCGACCCACGCGTCGGGAGT 82
```

```

RESULT 25
US-09-910-943-236
; Sequence 236, Application US/09910943
; Patent No. US20020081610A1
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Briuanlou, Ali
; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
; FILE REFERENCE: 7529/IG148US1
; CURRENT APPLICATION NUMBER: US/09/910,943
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 742
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 236
; LENGTH: 711
; TYPE: DNA
; ORGANISM: Xenopus laevis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(711)
; OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-236
```

```

Query Match          0.8%; Score 22; DB 9; Length 711;
Best Local Similarity 100.0%; Pred. No. 0.63;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 57 GTCGACCCACGCGTCGGGAGT 78
    |||||||
Db 48 GTCGACCCACGCGTCGGGAGT 69
```

```

RESULT 26
US-09-910-943-361
; Sequence 361, Application US/09910943
; Patent No. US20020081610A1
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Briuanlou, Ali
; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
; FILE REFERENCE: 7529/IG148US1
; CURRENT APPLICATION NUMBER: US/09/910,943
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 742
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 361
; LENGTH: 748
; TYPE: DNA
; ORGANISM: Xenopus laevis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(748)
; OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-361
```

```

Query Match          0.8%; Score 22; DB 9; Length 748;
Best Local Similarity 100.0%; Pred. No. 0.64;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 57 GTCGACCCACGCGTCGGGAGT 78
    |||||||
Db 49 GTCGACCCACGCGTCGGGAGT 70
```

```

RESULT 27
US-09-897-006-6/c
; Sequence 6, Application US/09897006
; Patent No. US20020106729A1
; GENERAL INFORMATION:
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Expression Vectors
; FILE REFERENCE: GALA-06415
; CURRENT APPLICATION NUMBER: US/09/897,006
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,851
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 5732
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-897-006-6
```

```

Query Match          0.8%; Score 22; DB 10; Length 5732;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 AATCTTTATTTATTCAGTGT 22
    |||||||
Db 5112 AATCTTTATTTATTCAGTGT 5091
```

```

RESULT 28
US-09-897-511A-6/c
; Sequence 6, Application US/09897511A
; Publication No. US20030092882A1
; GENERAL INFORMATION:
; APPLICANT: Bremel, Robert
; APPLICANT: Miller, Linda
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
```

FILE REFERENCE: GALA-06416
CURRENT APPLICATION NUMBER: US/09/897,511A
CURRENT FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 60/215,925
PRIOR FILING DATE: 2000-07-03
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.0
SEQ ID NO 6
LENGTH: 5732
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-09-897-511A-6

Query Match 0.8%; Score 22; DB 11; Length 5732;
Best Local Similarity 100.0%; Pred. No. 0.77;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGTT 22
|||||
Db 5112 AATCTTTATTTATCGATGTT 5091

RESULT 29
US-08-786-531B-4/c
Sequence 4, Application US/08786531B
Publication No. US20020015979A1
GENERAL INFORMATION:
APPLICANT: Link, Charles J.
APPLICANT: Levy, John P.
APPLICANT: Wang, Suming
APPLICANT: Seregina, Tatiana
TITLE OF INVENTION: Vehicles for Stable Transfer of Green
TITLE OF INVENTION: Fluorescent Protein Gene and Methods of Use for Same
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
STREET: 801 Grand Suite 3200
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,531B
FILING DATE: 21-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/010371
FILING DATE: 22-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Nebel, Heidi S.
REGISTRATION NUMBER: 37,719
REFERENCE/DOCKET NUMBER: hgtri
TELECOMMUNICATION INFORMATION:
TELEPHONE: 515-288-3667
TELEFAX: 515-288-1338
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 7352 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-786-531B-4

Query Match 0.8%; Score 22; DB 8; Length 7352;
Best Local Similarity 100.0%; Pred. No. 0.79;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGTT 22
|||||
Db 4371 AATCTTTATTTATCGATGTT 4350

RESULT 30
US-08-786-531B-1/c
Sequence 1, Application US/08786531B
Publication No. US20020015979A1
GENERAL INFORMATION:
APPLICANT: Link, Charles J.
APPLICANT: Levy, John P.
APPLICANT: Wang, Suming
APPLICANT: Seregina, Tatiana
TITLE OF INVENTION: Vehicles for Stable Transfer of Green
TITLE OF INVENTION: Fluorescent Protein Gene and Methods of Use for Same
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
STREET: 801 Grand Suite 3200
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/786,531B
FILING DATE: 21-JAN-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/010371
FILING DATE: 22-JAN-1996
ATTORNEY/AGENT INFORMATION:
NAME: Nebel, Heidi S.
REGISTRATION NUMBER: 37,719
REFERENCE/DOCKET NUMBER: hgtri
TELECOMMUNICATION INFORMATION:
TELEPHONE: 515-288-3667
TELEFAX: 515-288-1338
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 7353 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: circular
MOLECULE TYPE: other nucleic acid
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-786-531B-1

Query Match 0.8%; Score 22; DB 8; Length 7353;
Best Local Similarity 100.0%; Pred. No. 0.79;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGTT 22
|||||
Db 4372 AATCTTTATTTATCGATGTT 4351

RESULT 31
US-09-897-006-7/c
Sequence 7, Application US/09897006
Patent No. US20020106729A1
GENERAL INFORMATION:
APPLICANT: Bleck, Gregory


```

; TITLE OF INVENTION: Expression Vectors
; FILE REFERENCE: GALA-06415
; CURRENT APPLICATION NUMBER: US/09/897,006
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,851
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 9183
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-897-006-7

Query Match          0.8%; Score 22; DB 10; Length 9183;
Best Local Similarity 100.0%; Pred. No. 0.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AATCTTTATTTATCGATGTT 22
      |||||
Db      8563 AATCTTTATTTATCGATGTT 8542

RESULT 32
US-09-897-511A-7/c
; Sequence 7, Application US/09897511A
; Publication No. US20030092882A1
; GENERAL INFORMATION:
; APPLICANT: Bremel, Robert
; APPLICANT: Miller, Linda
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
; FILE REFERENCE: GALA-06415
; CURRENT APPLICATION NUMBER: US/09/897,511A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,925
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 9183
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-897-511A-7

Query Match          0.8%; Score 22; DB 11; Length 9183;
Best Local Similarity 100.0%; Pred. No. 0.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 AATCTTTATTTATCGATGTT 22
      |||||
Db      8563 AATCTTTATTTATCGATGTT 8542

RESULT 33
US-09-924-035A-528/c
; Sequence 528, Application US/09924035A
; Patent No. US20020142319A1
; GENERAL INFORMATION:
; APPLICANT: Grilach, Jim
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2011US
; CURRENT APPLICATION NUMBER: US/09/924,035A
; CURRENT FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,784
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: FastSeq for Windows Version 3.0
```

```

; SEQ ID NO 528
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(332)
; OTHER INFORMATION: n = A,T,C or G
US-09-924-035A-528

Query Match          0.8%; Score 21; DB 10; Length 332;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      57 GTGACCCACGCGTCCGGAG 77
      |||||
Db      322 GTGACCCACGCGTCCGGAG 302

RESULT 34
US-09-925-300-828
; Sequence 828, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 828
; LENGTH: 442
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (11)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (14)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (21)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (25)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (128)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (438)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-828

Query Match          0.8%; Score 21; DB 10; Length 442;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      57 GTGACCCACGCGTCCGGAG 77
      |||||
Db      56 GTGACCCACGCGTCCGGAG 76
```

RESULT 35
US-09-960-352-4373
; Sequence 4373, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4373
; LENGTH: 457
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 19-LIB3058-036-Q1-K1-E3
US-09-960-352-4373

Query Match 0.8%; Score 21; DB 10; Length 457;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 CGACCCACCGCTCCGGAGTA 79
|||||
Db 6 CGACCCACCGCTCCGGAGTA 26

RESULT 36
US-10-027-632-235696/c
; Sequence 235696, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; POLYMORPHISMS IN THE HUMAN GENOME
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 235696
; LENGTH: 690
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(690)
; OTHER INFORMATION: n = A,T,C or G
US-10-027-632-235696

Query Match 0.8%; Score 21; DB 13; Length 690;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 522 AAGACATCTGAAAGATGTG 542
|||||

Db 513 AAGACATCTGAAAGATGTG 493

RESULT 37
US-10-106-698-1858
; Sequence 1858, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA00591
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 1858
; LENGTH: 717
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (13)-(13)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (18)-(18)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (25)-(25)
; NAME/KEY: misc.feature
; LOCATION: (572)-(572)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-1858

Query Match 0.8%; Score 21; DB 14; Length 717;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGCACCCACCGCTCCGGAG 77
|||||
Db 76 GTGCACCCACCGCTCCGGAG 96

RESULT 38
US-09-910-943-67
; Sequence 67, Application US/09910943
; Patent No. US20020081610A1
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Briyanlou, Ali
; APPLICANT: Altman, Curtis
; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
; FILE REFERENCE: 7539/IG148US1
; CURRENT APPLICATION NUMBER: US/09/910,943
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 742
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67
; LENGTH: 725
; TYPE: DNA
; ORGANISM: Xenopus laevis
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(725)
; OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-67

Query Match 0.8%; Score 21; DB 9; Length 725;

Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTCGACCCACGCGTCGGGAG 77
|||||
Db 43 GTCGACCCACGCGTCGGGAG 63

RESULT 39

US-09-910-943-448
; Sequence 448, Application US/09910943
; Patent No. US20020081610A1
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Briyvanlou, Ali
; APPLICANT: Altman, Curtis
; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
; FILE REFERENCE: 7529/IG148US1
; CURRENT APPLICATION NUMBER: US/09/910,943
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 742
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 448
; LENGTH: 730
; TYPE: DNA
; ORGANISM: Xenopus laevis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)_(730)
; OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-448

Query Match 0.8%; Score 21; DB 9; Length 730;
Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTCGACCCACGCGTCGGGAG 77
|||||
Db 44 GTCGACCCACGCGTCGGGAG 64

RESULT 40

US-09-910-943-674
; Sequence 674, Application US/09910943
; Patent No. US20020081610A1
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Briyvanlou, Ali
; APPLICANT: Altman, Curtis
; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
; FILE REFERENCE: 7529/IG148US1
; CURRENT APPLICATION NUMBER: US/09/910,943
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 742
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 674
; LENGTH: 745
; TYPE: DNA
; ORGANISM: Xenopus laevis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)_(745)
; OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-674

Query Match 0.8%; Score 21; DB 9; Length 745;
Best Local Similarity 100.0%; Pred. No. 2.3;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTCGACCCACGCGTCGGGAG 77
|||||
Db 54 GTCGACCCACGCGTCGGGAG 74

RESULT 41

US-09-910-943-359
; Sequence 359, Application US/09910943
; Patent No. US20020081610A1
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Briyvanlou, Ali
; APPLICANT: Altman, Curtis
; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
; FILE REFERENCE: 7529/IG148US1
; CURRENT APPLICATION NUMBER: US/09/910,943
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 742
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 359
; LENGTH: 749
; TYPE: DNA
; ORGANISM: Xenopus laevis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)_(749)
; OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-359

Query Match 0.8%; Score 21; DB 9; Length 749;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTCGACCCACGCGTCGGGAG 77
|||||
Db 55 GTCGACCCACGCGTCGGGAG 75

RESULT 42

US-09-910-943-428
; Sequence 428, Application US/09910943
; Patent No. US20020081610A1
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Briyvanlou, Ali
; APPLICANT: Altman, Curtis
; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
; FILE REFERENCE: 7529/IG148US1
; CURRENT APPLICATION NUMBER: US/09/910,943
; CURRENT FILING DATE: 2001-07-23
; NUMBER OF SEQ ID NOS: 742
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 428
; LENGTH: 749
; TYPE: DNA
; ORGANISM: Xenopus laevis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)_(749)
; OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-428

Query Match 0.8%; Score 21; DB 9; Length 749;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTCGACCCACGCGTCGGGAG 77
|||||
Db 50 GTCGACCCACGCGTCGGGAG 70

RESULT 43

US-09-910-943-365
; Sequence 365, Application US/09910943
; Patent No. US20020081610A1
; GENERAL INFORMATION:
; APPLICANT: Hemmati-Briyvanlou, Ali
; APPLICANT: Altman, Curtis
; TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
; FILE REFERENCE: 7529/IG148US1
; CURRENT APPLICATION NUMBER: US/09/910,943

/ CURRENT FILING DATE: 2001-07-23
/ NUMBER OF SEQ ID NOS: 742
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 365
/ LENGTH: 750
/ TYPE: DNA
/ ORGANISM: Xenopus laevis
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)..(750)
/ OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-365

Query Match 0.8%; Score 21; DB 9; Length 750;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAG 77
DB 55 GTGACCCACGCGTCCGGAG 75

RESULT 44
US-09-910-943-504
/ Sequence 504, Application US/09910943
/ Patent No. US20020081610A1
/ GENERAL INFORMATION:
/ APPLICANT: Hemmati-Briyvanlou, Ali
/ APPLICANT: Altman, Curtis
/ TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
/ FILE REFERENCE: 7529/1G148US1
/ CURRENT APPLICATION NUMBER: US/09/910,943
/ CURRENT FILING DATE: 2001-07-23
/ NUMBER OF SEQ ID NOS: 742
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 504
/ LENGTH: 750
/ TYPE: DNA
/ ORGANISM: Xenopus laevis
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)..(750)
/ OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-504

Query Match 0.8%; Score 21; DB 9; Length 750;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAG 77
DB 49 GTGACCCACGCGTCCGGAG 69

RESULT 45
US-09-910-943-383
/ Sequence 383, Application US/09910943
/ Patent No. US20020081610A1
/ GENERAL INFORMATION:
/ APPLICANT: Hemmati-Briyvanlou, Ali
/ APPLICANT: Altman, Curtis
/ TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
/ FILE REFERENCE: 7529/1G148US1
/ CURRENT APPLICATION NUMBER: US/09/910,943
/ CURRENT FILING DATE: 2001-07-23
/ NUMBER OF SEQ ID NOS: 742
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 383
/ LENGTH: 781
/ TYPE: DNA
/ ORGANISM: Xenopus laevis
/ FEATURE:
/ NAME/KEY: misc_feature

/ LOCATION: (1)..(781)
/ OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-383

Query Match 0.8%; Score 21; DB 9; Length 781;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAG 77
DB 43 GTGACCCACGCGTCCGGAG 63

RESULT 46
US-09-910-943-609
/ Sequence 609, Application US/09910943
/ Patent No. US20020081610A1
/ GENERAL INFORMATION:
/ APPLICANT: Hemmati-Briyvanlou, Ali
/ APPLICANT: Altman, Curtis
/ TITLE OF INVENTION: Assays and Materials for Embryonic Gene Expression
/ FILE REFERENCE: 7529/1G148US1
/ CURRENT APPLICATION NUMBER: US/09/910,943
/ CURRENT FILING DATE: 2001-07-23
/ NUMBER OF SEQ ID NOS: 742
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 609
/ LENGTH: 784
/ TYPE: DNA
/ ORGANISM: Xenopus laevis
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)..(784)
/ OTHER INFORMATION: n may be a or g or c or t/u
US-09-910-943-609

Query Match 0.8%; Score 21; DB 9; Length 784;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAG 77
DB 68 GTGACCCACGCGTCCGGAG 88

RESULT 47
US-10-198-846-14063
/ Sequence 14063, Application US/10198846
/ Publication No. US2003009974A1
/ GENERAL INFORMATION:
/ APPLICANT: Lillie, James
/ APPLICANT: Xu, Yongyao
/ APPLICANT: Wang, Youzhen
/ APPLICANT: Steinmann, Kathleen
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
/ TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
/ FILE REFERENCE: MRI-049
/ CURRENT APPLICATION NUMBER: US/10/198,846
/ CURRENT FILING DATE: 2002-07-18
/ PRIOR APPLICATION NUMBER: 60/306,220
/ PRIOR FILING DATE: 2001-07-18
/ NUMBER OF SEQ ID NOS: 14084
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 14063
/ LENGTH: 883
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-198-846-14063

Query Match 0.8%; Score 21; DB 14; Length 883;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAG 77
DB 4 GTGACCCACGCGTCCGGAG 24

RESULT 48

US-10-037-270-33
; Sequence 33, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungting
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: Tillinghast, John
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/10/037,270
CURRENT FILING DATE: 2002-01-04
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 33
LENGTH: 1181
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (173)..(874)
US-10-037-270-33

Query Match

Best Local Similarity 0.8%; Score 21; DB 14; Length 1181;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAG 77
DB 59 GTGACCCACGCGTCCGGAG 79

RESULT 49

US-10-106-698-1342
; Sequence 1342, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:

APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
FILE REFERENCE: PA005P1
CURRENT APPLICATION NUMBER: US/10/106,698
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: PCT/US00/26524
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280
PRIOR FILING DATE: 1999-11-03

NUMBER OF SEQ ID NOS: 8564
SOFTWARE: PatentIn Ver. 3.0
SEQ ID NO 1342
LENGTH: 1280
TYPE: DNA
ORGANISM: Homo sapiens

FEATURE:
NAME/KEY: misc feature
LOCATION: (29)..(29)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (47)..(47)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (83)..(83)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (121)..(121)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (133)..(133)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (154)..(154)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc feature
LOCATION: (1166)..(1166)
OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-1342

Query Match

Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAG 77
DB 243 GTGACCCACGCGTCCGGAG 263

RESULT 50

US-09-925-776-1
; Sequence 1, Application US/09925776
; Patent No. US20020038007A1
; GENERAL INFORMATION:

APPLICANT: AMES, ROBERT S., JR.
APPLICANT: SARAU, HENRY M.
APPLICANT: FOLEY, JAMES J.
APPLICANT: BERGSMAN, DEREK J.
APPLICANT: ELLIS, CATHERINE E.
TITLE OF INVENTION: A METHOD OF FINDING AGONIST AND
TITLE OF INVENTION: ANTAGONIST TO HUMAN 11CB SPLICED VARIANT
FILE REFERENCE: GP-50003-D2
CURRENT APPLICATION NUMBER: US/09/925,776
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 60/032,763
PRIOR FILING DATE: 1996-12-11
PRIOR APPLICATION NUMBER: 06/984,288
PRIOR FILING DATE: 1997-12-03
PRIOR APPLICATION NUMBER: 60/073,747
PRIOR FILING DATE: 1998-02-05
PRIOR APPLICATION NUMBER: 09/060,504
PRIOR FILING DATE: 1998-04-15
NUMBER OF SEQ ID NOS: 7
SOFTWARE: FastSeq For Windows Version 3.0
SEQ ID NO 1
LENGTH: 1385
TYPE: DNA
ORGANISM: HOMO SAPIENS
US-09-925-776-1

Query Match 0.8%; Score 21; DB 9; Length 1385;
Best Local Similarity 100.0%; Pred. No. 2.4;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCCGGAG 77
|||||
Db 50 GTGACCCACGCGTCCGGAG 70

RESULT 51
US-10-119-926-44
; Sequence 44, Application US/10119926
; Publication No. US20030104413A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhao, Qing A.
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wehrman, Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Xu, Chongjun
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Dirmnac, Radoje T.
; TITLE OF INVENTION: No. US20030104413A1 Nucleic Acids and
; FILE REFERENCE: Polypeptides
; FILE REFERENCE: 789CIP2BCON
; CURRENT APPLICATION NUMBER: US/10/119,926
; CURRENT FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 09/574,454
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: 09/519,705
; PRIOR FILING DATE: 2000-03-07
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 44
; LENGTH: 1394
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (415)..(1227)
US-10-119-926-44

Query Match 0.8%; Score 21; DB 14; Length 1394;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCCGGAG 77
|||||
Db 33 GTGACCCACGCGTCCGGAG 53

RESULT 52
US-10-037-270-664
; Sequence 664, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunting
; APPLICANT: Wang, Dunrui

; APPLICANT: Wang, Zhiwei
; APPLICANT: Tillinghast, John
; APPLICANT: Dirmnac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1 Nucleic Acids and
; FILE REFERENCE: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 664
; LENGTH: 1572
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (330)..(704)
US-10-037-270-664

Query Match 0.8%; Score 21; DB 14; Length 1572;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCCGGAG 77
|||||
Db 280 GTGACCCACGCGTCCGGAG 300

RESULT 53
US-09-838-573-1
; Sequence 1, Application US/09838573
; Patent No. US20020034783A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel
; TITLE OF INVENTION: 39228, A NOVEL HUMAN ALCOHOL
; FILE REFERENCE: MNT-143
; CURRENT APPLICATION NUMBER: US/09/838,573
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/197,747
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1808
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (285)..(1418)
US-09-838-573-1

Query Match 0.8%; Score 21; DB 9; Length 1808;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCCGGAG 77
|||||
Db 4 GTGACCCACGCGTCCGGAG 24

RESULT 54
US-10-160-501-1
; Sequence 1, Application US/10160501
; Publication No. US20030059919A1
; GENERAL INFORMATION:
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Rudolph-Owen, Laura A.
; APPLICANT: Kapeller-Libermann, Rosana

```
; TITLE OF INVENTION: NOVEL HUMAN 39228, 21956, 25856, 22244, 8701, 32263,
; TITLE OF INVENTION: 50250, 55158, 47765, 62088, 50566, AND 48118
; TITLE OF INVENTION: MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-250
; CURRENT APPLICATION NUMBER: US/10/160,501
; CURRENT FILING DATE: 2002-05-30
; PRIOR APPLICATION NUMBER: US 09/838,573
; PRIOR FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/197,747
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 09/870,133
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,649
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/870,130
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,640
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/862,535
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/205,961
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/870,383
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,506
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/860,821
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,449
; PRIOR FILING DATE: 2000-05-19
; PRIOR APPLICATION NUMBER: US 09/870,110
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: US 60/207,650
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/907,509
; PRIOR FILING DATE: 2001-07-16
; PRIOR APPLICATION NUMBER: US 60/218,385
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 09/945,327
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/229,425
; PRIOR FILING DATE: 2000-08-31
; PRIOR APPLICATION NUMBER: US 60/318,581
; PRIOR FILING DATE: 2001-09-10
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: FastSeq Version 4.0
; SEQ ID NO 1
; LENGTH: 1808
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (285)...(1418)
; US-10-160-501-1

Query Match          0.8%; Score 21; DB 14; Length 1808;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 57 GTCGACCCAGCGTCGGGAG 77
Db 4 GTCGACCCAGCGTCGGGAG 24
```

```
RESULT 55
US-09-925-302-192
; Sequence 192, Application US/09925302
; Patent No. US20020044941A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA104
; CURRENT APPLICATION NUMBER: US/09/925,302
```

```
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05918
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 896
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 192
; LENGTH: 2570
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2561)
; OTHER INFORMATION: n equals a,t,g, or c
; US-09-925-302-192
```

```
Query Match          0.8%; Score 21; DB 9; Length 2570;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 58 TCGACCCAGCGTCGGGAGT 78
Db 1 TCGACCCAGCGTCGGGAGT 21
```

```
RESULT 56
US-09-796-753-89
; Sequence 89, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT APPLICATION NUMBER: US/09/796,753
; CURRENT FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,687
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 09/345,464
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/365,164
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/399,723
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 09/409,634
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/471,179
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/474,072
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
```

```
/ PRIOR FILING DATE: 2000-05-14
/ PRIOR APPLICATION NUMBER: 09/597,993
/ PRIOR FILING DATE: 2000-06-19
/ PRIOR APPLICATION NUMBER: 09/599,596
/ PRIOR FILING DATE: 2000-06-22
/ PRIOR APPLICATION NUMBER: 09/630,334
/ PRIOR FILING DATE: 2000-07-31
/ PRIOR APPLICATION NUMBER: 09/606,565
/ PRIOR FILING DATE: 2000-06-29
/ PRIOR APPLICATION NUMBER: 09/606,317
/ PRIOR FILING DATE: 2000-06-29
/ PRIOR APPLICATION NUMBER: 09/665,666
/ PRIOR FILING DATE: 2000-09-20
/ PRIOR APPLICATION NUMBER: 09/677,751
/ PRIOR FILING DATE: 2000-09-30
/ NUMBER OF SEQ ID NOS: 162
/ SEQ ID NO 89
/ LENGTH: 2810
/ TYPE: DNA
/ ORGANISM: Mus musculus
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (187)..(1527)
US-09-796-753-89

Query Match          0.8%; Score 21; DB 11; Length 2810;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCGGGAG 77
Db 1 GTGACCCACGCGTCGGGAG 21

RESULT 57
US-09-782-980-67
/ Sequence 67, Application US/09782980
/ Patent No. US20020072089A1
/ GENERAL INFORMATION:
/ APPLICANT: Khodadoust, Mehran M.
/ APPLICANT: Macbeth, Kyle J.
/ APPLICANT: Busfield, Samantha J.
/ APPLICANT: McCarthy, Sean A.
/ APPLICANT: Holtzman, Douglas A.
/ APPLICANT: Gu, Wei
/ APPLICANT: White, David
/ APPLICANT: Pan, Yang
/ TITLE OF INVENTION: NOVEL ITALY, LOR-2, STRIPE, TRASH, BDSF, LBSG, AND
/ TITLE OF INVENTION: STWST PROTEIN AND NUCLEIC ACID MOLECULES AND USES
/ FILE REFERENCE: MNI-121CP
/ CURRENT APPLICATION NUMBER: US/09/782,980
/ PRIOR APPLICATION NUMBER: PCT/US00/02125
/ PRIOR FILING DATE: 2001-02-13
/ PRIOR APPLICATION NUMBER: 09/448,076
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: 09/276,400
/ PRIOR FILING DATE: 1999-03-25
/ PRIOR APPLICATION NUMBER: 60/117,580
/ PRIOR FILING DATE: 1999-01-27
/ PRIOR APPLICATION NUMBER: 09/014,195
/ PRIOR FILING DATE: 1998-01-27
/ PRIOR APPLICATION NUMBER: 09/014,348
/ PRIOR FILING DATE: 1998-01-27
/ PRIOR APPLICATION NUMBER: 09/086,892
/ PRIOR FILING DATE: 1998-05-29
/ PRIOR APPLICATION NUMBER: 09/296,208
/ PRIOR FILING DATE: 1999-04-21
/ PRIOR APPLICATION NUMBER: 09/063,950
/ PRIOR FILING DATE: 1998-04-21
/ PRIOR APPLICATION NUMBER: 09/561,381
/ PRIOR FILING DATE: 2000-04-28
```

```
/ PRIOR APPLICATION NUMBER: 09/561,810
/ PRIOR FILING DATE: 2000-04-28
/ PRIOR APPLICATION NUMBER: 09/087,121
/ PRIOR FILING DATE: 1998-05-29
/ PRIOR APPLICATION NUMBER: 09/672,721
/ PRIOR FILING DATE: 2000-09-28
/ PRIOR APPLICATION NUMBER: 09/049,799
/ PRIOR FILING DATE: 1998-03-27
/ NUMBER OF SEQ ID NOS: 176
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 67
/ LENGTH: 2815
/ TYPE: DNA
/ ORGANISM: Mus musculus
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (197)..(2215)
US-09-782-980-67

Query Match          0.8%; Score 21; DB 9; Length 2815;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCGGGAG 77
Db 1 GTGACCCACGCGTCGGGAG 21

RESULT 58
US-09-897-006-4/c
/ Sequence 4, Application US/09897006
/ Patent No. US20020106729A1
/ GENERAL INFORMATION:
/ APPLICANT: Bleck, Gregory
/ TITLE OF INVENTION: Expression Vectors
/ FILE REFERENCE: GALA-06415
/ CURRENT APPLICATION NUMBER: US/09/897,006
/ PRIOR FILING DATE: 2001-06-29
/ PRIOR APPLICATION NUMBER: 60/215,851
/ PRIOR FILING DATE: 2000-07-03
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 4
/ LENGTH: 4207
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Synthetic
US-09-897-006-4

Query Match          0.8%; Score 21; DB 10; Length 4207;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGT 21
Db 3587 AATCTTTATTTATCGATGT 3567

RESULT 59
US-09-897-511A-4/c
/ Sequence 4, Application US/09897511A
/ Publication No. US20030092882A1
/ GENERAL INFORMATION:
/ APPLICANT: Bremel, Robert
/ APPLICANT: Miller, Linda
/ APPLICANT: Bleck, Gregory
/ TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
/ FILE REFERENCE: GALA-06416
/ CURRENT APPLICATION NUMBER: US/09/897,511A
/ PRIOR FILING DATE: 2001-06-29
/ PRIOR APPLICATION NUMBER: 60/215,925
/ PRIOR FILING DATE: 2000-07-03
```


NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.0
SEQ ID NO 4
LENGTH: 4207
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-09-897-511A-4

Query Match 0.8%; Score 21; DB 11; Length 4207;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGT 21
|||||
Db 3587 AATCTTTATTTATCGATGT 3567

RESULT 60
US-09-897-006-5/c
Sequence 5, Application US/09897006
Patent No. US20020106729A1
GENERAL INFORMATION:
APPLICANT: Bleck, Gregory
TITLE OF INVENTION: Expression Vectors
FILE REFERENCE: GALA-06415
CURRENT APPLICATION NUMBER: US/09/897,006
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 60/215,851
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 4210
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-09-897-006-5

Query Match 0.8%; Score 21; DB 10; Length 4210;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGT 21
|||||
Db 3595 AATCTTTATTTATCGATGT 3575

RESULT 61
US-09-897-511A-5/c
Sequence 5, Application US/09897511A
Publication No. US20030092882A1
GENERAL INFORMATION:
APPLICANT: Bremel, Robert
APPLICANT: Miller, Linda
APPLICANT: Bleck, Gregory
TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
FILE REFERENCE: GALA-06416
CURRENT APPLICATION NUMBER: US/09/897,511A
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 60/215,925
PRIOR FILING DATE: 2000-07-03
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.0
SEQ ID NO 5
LENGTH: 4210
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-09-897-511A-5

Query Match 0.8%; Score 21; DB 11; Length 4210;
Best Local Similarity 100.0%; Pred. No. 2.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGT 21
|||||
Db 3595 AATCTTTATTTATCGATGT 3575

RESULT 62
US-09-897-006-13/c
Sequence 13, Application US/09897006
Patent No. US20020106729A1
GENERAL INFORMATION:
APPLICANT: Bleck, Gregory
TITLE OF INVENTION: Expression Vectors
FILE REFERENCE: GALA-06415
CURRENT APPLICATION NUMBER: US/09/897,006
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 60/215,851
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.0
SEQ ID NO 13
LENGTH: 6255
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-09-897-006-13

Query Match 0.8%; Score 21; DB 10; Length 6255;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGT 21
|||||
Db 5635 AATCTTTATTTATCGATGT 5615

RESULT 63
US-09-897-511A-13/c
Sequence 13, Application US/09897511A
Publication No. US20030092882A1
GENERAL INFORMATION:
APPLICANT: Bremel, Robert
APPLICANT: Miller, Linda
APPLICANT: Bleck, Gregory
TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
FILE REFERENCE: GALA-06416
CURRENT APPLICATION NUMBER: US/09/897,511A
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 60/215,925
PRIOR FILING DATE: 2000-07-03
NUMBER OF SEQ ID NOS: 36
SOFTWARE: PatentIn version 3.0
SEQ ID NO 13
LENGTH: 6255
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
US-09-897-511A-13

Query Match 0.8%; Score 21; DB 11; Length 6255;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGT 21
|||||
Db 5635 AATCTTTATTTATCGATGT 5615

RESULT 64
US-09-897-006-34/c
; Sequence 34, Application US/09897006
; Patent No. US20020106729A1
; GENERAL INFORMATION:
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Expression Vectors
; FILE REFERENCE: GALA-06415
; CURRENT APPLICATION NUMBER: US/09/897,006
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,851
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 9511
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-897-006-34

Query Match 0.8%; Score 21; DB 10; Length 9511;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATG 21
|||||
Db 6526 AATCTTTATTTATCGATG 6506

RESULT 65
US-09-897-511A-34/c
; Sequence 34, Application US/09897511A
; Publication No. US20030092862A1
; GENERAL INFORMATION:
; APPLICANT: Bremel, Robert
; APPLICANT: Miller, Linda
; APPLICANT: Bleck, Gregory
; TITLE OF INVENTION: Host Cells Containing Multiple Integrating Vectors
; FILE REFERENCE: GALA-06416
; CURRENT APPLICATION NUMBER: US/09/897,511A
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 60/215,925
; PRIOR FILING DATE: 2000-07-03
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 34
; LENGTH: 9511
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-897-511A-34

Query Match 0.8%; Score 21; DB 11; Length 9511;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATG 21
|||||
Db 6526 AATCTTTATTTATCGATG 6506

RESULT 66
US-10-106-698-4090
; Sequence 4090, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005Pl

; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1998-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 4090
; LENGTH: 100
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (49)..(49)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (62)..(62)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (74)..(74)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (88)..(88)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc.feature
; LOCATION: (93)..(93)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-4090

Query Match 0.7%; Score 20; DB 14; Length 100;
Best Local Similarity 100.0%; Pred. No. 6.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCGGGA 76
|||||
Db 24 GTGACCCACGCGTCGGGA 43

RESULT 67
US-09-764-891-522
; Sequence 522, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 522
; LENGTH: 122
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (121)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-891-522

Query Match 0.7%; Score 20; DB 11; Length 122;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCGGGA 76
|||||
Db 31 GTGACCCACGCGTCGGGA 50

```
RESULT 68
US-10-106-698-4018
; Sequence 4018, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 4018
; LENGTH: 142
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (8)..(9)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (18)..(18)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (27)..(27)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (31)..(31)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (107)..(107)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (125)..(125)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (128)..(128)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (135)..(135)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-4018

Query Match      0.7%; Score 20; DB 14; Length 142;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTCGACCCACGCGTCCGGGA 76
      |||||||
DB      69 GTCGACCCACGCGTCCGGGA 88

RESULT 69
US-10-106-698-2365
; Sequence 2365, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
```

```
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 2365
; LENGTH: 197
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (15)..(15)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (21)..(22)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (27)..(27)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (41)..(41)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (192)..(192)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (194)..(197)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-2365

Query Match      0.7%; Score 20; DB 14; Length 197;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTCGACCCACGCGTCCGGGA 76
      |||||||
DB      125 GTCGACCCACGCGTCCGGGA 144

RESULT 70
US-09-918-995-8601
; Sequence 8601, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8601
; LENGTH: 327
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(327)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-8601

Query Match      0.7%; Score 20; DB 11; Length 327;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTCGACCCACGCGTCCGGGA 76
      |||||||
DB      32 GTCGACCCACGCGTCCGGGA 51

RESULT 71
US-09-986-480-113
```

```
/ Sequence 113, Application US/09986480
/ Publication No. US20030027999A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: 143 Human Secreted Proteins
/ FILE REFERENCE: PS500P1
/ CURRENT APPLICATION NUMBER: US/09/986,480
/ PRIOR FILING DATE: 2001-11-08
/ PRIOR APPLICATION NUMBER: PCT/US00/12788
/ PRIOR FILING DATE: 2000-05-11
/ PRIOR APPLICATION NUMBER: US 60/134,068
/ PRIOR FILING DATE: 1999-05-13
/ NUMBER OF SEQ ID NOS: 456
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 113
/ LENGTH: 355
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-986-480-113

Query Match
Best Local Similarity 100.0%; Score 20; DB 11; Length 355;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGGCTCCGGGA 76
DB 5 GTGACCCACGGCTCCGGGA 24

RESULT 72
US-10-198-846-10302/c
/ Sequence 10302, Application US/10198846
/ Publication No. US20030099744A1
/ GENERAL INFORMATION:
/ APPLICANT: Lillie, James
/ APPLICANT: Xu, Yongyao
/ APPLICANT: Wang, Youzhen
/ APPLICANT: Steinhilber, Kathleen
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
/ TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
/ TITLE OF INVENTION: THERAPY OF BREAST CANCER
/ FILE REFERENCE: MRI-049
/ CURRENT APPLICATION NUMBER: US/10/198,846
/ CURRENT FILING DATE: 2002-07-18
/ PRIOR APPLICATION NUMBER: 60/306,220
/ PRIOR FILING DATE: 2001-07-18
/ NUMBER OF SEQ ID NOS: 14084
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 10302
/ LENGTH: 367
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-198-846-10302

Query Match
Best Local Similarity 100.0%; Score 20; DB 14; Length 367;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGGCTCCGGGA 76
DB 362 GTGACCCACGGCTCCGGGA 343

RESULT 73
US-09-925-300-802/c
/ Sequence 802, Application US/09925300
/ Patent No. US20020151681A1
/ GENERAL INFORMATION:
/ APPLICANT: Craig Rosen,
/ APPLICANT: Steve Ruben,
/ TITLE OF INVENTION: Nucleic Acids, Proteins and Antipodites
/ FILE REFERENCE: PA101
/ CURRENT APPLICATION NUMBER: US/09/925,300
```

```
/ CURRENT FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: PCT/US00/05988
/ PRIOR FILING DATE: 2000-03-08
/ PRIOR APPLICATION NUMBER: 60/124,270
/ PRIOR FILING DATE: 1999-03-12
/ NUMBER OF SEQ ID NOS: 1890
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 802
/ LENGTH: 402
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (147)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (149)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (310)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (322)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (344)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (363)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (383)
/ OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-802

Query Match
Best Local Similarity 100.0%; Score 20; DB 10; Length 402;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGGCTCCGGGA 76
DB 41 GTGACCCACGGCTCCGGGA 22

RESULT 74
US-10-106-698-2172
/ Sequence 2172, Application US/10106698
/ Publication No. US20030109690A1
/ GENERAL INFORMATION:
/ APPLICANT: Ruben et al.
/ TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
/ FILE REFERENCE: PA005P1
/ CURRENT APPLICATION NUMBER: US/10/106,698
/ CURRENT FILING DATE: 2002-03-27
/ PRIOR APPLICATION NUMBER: PCT/US00/26524
/ PRIOR FILING DATE: 2000-09-28
/ PRIOR APPLICATION NUMBER: US 60/157,137
/ PRIOR FILING DATE: 1999-09-29
/ PRIOR APPLICATION NUMBER: US 60/163,280
/ PRIOR FILING DATE: 1999-11-03
/ NUMBER OF SEQ ID NOS: 8564
/ SOFTWARE: PatentIn Ver. 3.0
/ SEQ ID NO 2172
/ LENGTH: 405
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (24)-(24)
/ OTHER INFORMATION: n equals a,t,g, or c
/ NAME/KEY: misc feature
/ LOCATION: (357)-(357)
/ OTHER INFORMATION: n equals a,t,g, or c
```

NAME/KEY: misc.feature
LOCATION: (385)..(385)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: misc.feature
LOCATION: (394)..(396)
OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-2172

Query Match 0.7%; Score 20; DB 14; Length 405;
Best Local Similarity 100.0%; Pred. No. 7.6; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTCGACCCAGCGCTCCGGGA 76
Db 72 GTCGACCCAGCGCTCCGGGA 91

RESULT 75
US-09-960-352-13318
Sequence 13318, Application US/09960352
Patent No. US20020137139A1
GENERAL INFORMATION:
APPLICANT: Warren, Wesley C.
APPLICANT: Tao, Nengbing
APPLICANT: Byatt, John C.
APPLICANT: Mathialagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 16511.006/37-21(10298)C
CURRENT APPLICATION NUMBER: US/09/960,352
CURRENT FILING DATE: 2001-09-24
NUMBER OF SEQ ID NOS: 15112
SEQ ID NO 13318
LENGTH: 430
TYPE: DNA
ORGANISM: Bos taurus
OTHER INFORMATION: Clone ID: 57-LIB3057-004-Q1-K1-G2
US-09-960-352-13318

Query Match 0.7%; Score 20; DB 10; Length 430;
Best Local Similarity 100.0%; Pred. No. 7.6; Mismatches 0; Indels 0; Gaps 0;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 61 ACCCAGCGCTCCGGGAGTAG 80
Db 1 ACCCAGCGCTCCGGGAGTAG 20

Search completed: August 1, 2003, 02:13:41
Job time : 602 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: July 31, 2003, 19:37:11 ; Search time 7225 Seconds

(without alignments)
12286.241 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709
Sequence: 1 aactcttatttcgacg.....eggtcttttccataacc 2709

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 32997241 seqs, 1638322548 residues

Word size : 0

Total number of hits satisfying chosen parameters: 65994482

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 100 summaries

Database :

Pending Patents NA Main:*

1: /cgn2_6/ptodata/1/pna/PCRTUS_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/PCRTUS_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US08_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US092_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US092B_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US093A_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US093B_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US098A_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US098B_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US098C_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US098D_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US099E_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US099F_COMB.seq.*

44: /cgn2_6/ptodata/1/pna/US100A_COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US100B_COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US101A_COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US101B_COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US102A_COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US102B_COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US103A_COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US103B_COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US104A_COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US104B_COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US6000_COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US6001_COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US6002_COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US6003_COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US6004_COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US6005_COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US6006_COMB.seq.*
61: /cgn2_6/ptodata/1/pna/US6007_COMB.seq.*
62: /cgn2_6/ptodata/1/pna/US6008_COMB.seq.*
63: /cgn2_6/ptodata/1/pna/US6009_COMB.seq.*
64: /cgn2_6/ptodata/1/pna/US6010_COMB.seq.*
65: /cgn2_6/ptodata/1/pna/US6011_COMB.seq.*
66: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
67: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
68: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
69: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
70: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
71: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
72: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
73: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
74: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
75: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
76: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
77: /cgn2_6/ptodata/1/pna/US6023A_COMB.seq.*
78: /cgn2_6/ptodata/1/pna/US6023B_COMB.seq.*
79: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
80: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
81: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
82: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
83: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*
84: /cgn2_6/ptodata/1/pna/US6029_COMB.seq.*
85: /cgn2_6/ptodata/1/pna/US6030_COMB.seq.*
86: /cgn2_6/ptodata/1/pna/US6031_COMB.seq.*
87: /cgn2_6/ptodata/1/pna/US6032_COMB.seq.*
88: /cgn2_6/ptodata/1/pna/US6033_COMB.seq.*
89: /cgn2_6/ptodata/1/pna/US6034_COMB.seq.*
90: /cgn2_6/ptodata/1/pna/US6035_COMB.seq.*
91: /cgn2_6/ptodata/1/pna/US6036_COMB.seq.*
92: /cgn2_6/ptodata/1/pna/US6037_COMB.seq.*
93: /cgn2_6/ptodata/1/pna/US6038_COMB.seq.*
94: /cgn2_6/ptodata/1/pna/US6039_COMB.seq.*
95: /cgn2_6/ptodata/1/pna/US6040_COMB.seq.*
96: /cgn2_6/ptodata/1/pna/US6041_COMB.seq.*
97: /cgn2_6/ptodata/1/pna/US6042_COMB.seq.*
98: /cgn2_6/ptodata/1/pna/US6043_COMB.seq.*
99: /cgn2_6/ptodata/1/pna/US6044_COMB.seq.*
100: /cgn2_6/ptodata/1/pna/US6045_COMB.seq.*
101: /cgn2_6/ptodata/1/pna/US6046_COMB.seq.*
102: /cgn2_6/ptodata/1/pna/US6047_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2709	100.0	2709	US-09-774-490-1	Sequence 1, Appl
2	2517	92.9	2848	US-10-247-671-83	Sequence 83, Appl
3	2517	92.9	2848	US-60-172-373-8935	Sequence 8935, Ap
4	2517	92.9	2848	US-60-323-784-83	Sequence 83, Appl

5	2508	92.6	2530	1	PCT-US02-07826-283	Sequence 283, App
6	2508	92.6	2530	15	US-09-053-375B-1350	Sequence 1350, Ap
7	2508	92.6	2530	45	US-10-097-340-283	Sequence 283, App
8	2508	92.6	2530	49	US-10-262-538-9	Sequence 9, Appl
9	2508	92.6	2601	9	US-08-488-935-1	Sequence 1, Appl
10	2508	92.6	2601	9	US-08-489-057-1	Sequence 1, Appl
11	2508	92.6	2601	45	US-10-067-632-53	Sequence 53, Appl
12	2226	82.2	2857	47	US-10-170-235-19890	Sequence 19890, Ap
13	1388	51.2	2256	28	US-09-649-164-8352	Sequence 8352, Ap
14	1388	51.2	2256	28	US-09-652-816-7308	Sequence 7308, Ap
15	1388	51.2	2256	28	US-09-652-918-7516	Sequence 7516, Ap
16	817	30.2	944	32	US-09-724-676-18078	Sequence 18078, Ap
17	817	30.2	944	32	US-09-724-676A-18078	Sequence 18078, A
18	405	15.0	456	36	US-09-864-761-30392	Sequence 30392, A
19	405	15.0	456	48	US-10-203-135-19786	Sequence 19786, A
20	405	15.0	456	48	US-10-203-135-20203	Sequence 20203, A
21	379	14.0	379	17	US-09-205-070-45112	Sequence 45112, A
22	379	14.0	379	19	US-09-321-214-10943	Sequence 10943, A
23	379	14.0	379	19	US-09-340-623-45112	Sequence 45112, A
24	379	14.0	379	22	US-09-516-335-10943	Sequence 10943, A
25	379	14.0	379	32	US-09-733-811-10943	Sequence 10943, A
26	379	14.0	379	32	US-09-733-811A-10943	Sequence 10943, A
27	379	14.0	379	37	US-09-898-888-45112	Sequence 45112, A
28	379	14.0	379	37	US-09-898-888A-45112	Sequence 45112, A
29	379	14.0	379	42	US-09-975-640-10943	Sequence 10943, A
30	379	14.0	379	42	US-09-975-640A-10943	Sequence 10943, A
31	370	13.7	497	23	US-09-534-856-4	Sequence 4, Appl
32	303	11.2	354	1	PCT-US02-11475A-736	Sequence 736, App
33	303	11.2	354	26	US-09-609-448A-736	Sequence 736, App
34	303	11.2	354	28	US-09-649-811-736	Sequence 736, App
35	303	11.2	354	35	US-09-833-263-736	Sequence 736, App
36	303	11.2	354	39	US-09-922-217-736	Sequence 736, App
37	303	11.2	354	44	US-10-025-380-736	Sequence 736, App
38	292	10.8	496	36	US-09-864-761-14444	Sequence 14444, A
39	292	10.8	496	48	US-10-203-135-7281	Sequence 7281, Ap
40	292	10.8	496	48	US-10-203-136-7203	Sequence 7203, Ap
41	292	10.8	496	48	US-10-203-136-7205	Sequence 7205, A
42	256	9.4	427	20	US-09-399-720-14058	Sequence 14058, A
43	256	9.4	427	39	US-09-921-378-14058	Sequence 14058, A
44	256	9.4	427	21	US-09-442-385-770	Sequence 770, App
45	228	8.4	268	34	US-09-824-130-5891	Sequence 5891, Ap
46	169	6.2	636	70	US-60-160-203-79	Sequence 79, Appl
47	161	6.2	636	70	US-60-169-840-156	Sequence 156, App
48	161	5.9	164	44	US-10-029-386-18184	Sequence 18184, A
49	161	5.9	591	44	US-10-029-386-4484	Sequence 4484, Ap
50	160	5.9	172	1	PCT-US01-00663-20125	Sequence 20125, A
51	160	5.9	172	2	PCT-US01-00663-20125	Sequence 20125, A
52	160	5.9	172	36	US-09-864-761-27943	Sequence 27943, A
53	160	5.9	172	47	US-10-182-993-19495	Sequence 19495, A
54	160	5.9	172	48	US-10-203-134-20052	Sequence 20052, A
55	160	5.9	172	48	US-10-203-135-19754	Sequence 19754, A
56	160	5.9	172	48	US-10-203-136-20171	Sequence 20171, A
57	160	5.9	172	48	US-10-203-137-19125	Sequence 20125, A
58	160	5.9	191	44	US-10-203-139-19510	Sequence 19510, A
59	160	5.9	191	44	US-10-029-386-16491	Sequence 16491, A
60	160	5.9	484	1	PCT-US01-00663-7072	Sequence 7072, Ap
61	160	5.9	484	2	PCT-US01-00663-7072	Sequence 7072, Ap
62	160	5.9	484	36	US-09-864-761-11363	Sequence 11363, A
63	160	5.9	484	47	US-10-182-993-6833	Sequence 6833, Ap
64	160	5.9	484	48	US-10-203-134-7118	Sequence 7118, Ap
65	160	5.9	484	48	US-10-203-135-7342	Sequence 7342, Ap
66	160	5.9	484	48	US-10-203-136-7262	Sequence 7262, Ap
67	160	5.9	484	48	US-10-203-137-7012	Sequence 7012, Ap
68	160	5.9	526	44	US-10-029-386-2791	Sequence 2791, Ap
69	144	5.3	446	36	US-09-864-761-11810	Sequence 11810, A
70	144	5.3	446	47	US-10-182-993-7713	Sequence 7713, Ap
71	144	5.3	446	48	US-10-203-135-8417	Sequence 8417, Ap
72	144	5.3	446	48	US-10-203-136-8331	Sequence 8331, Ap
73	144	5.3	574	70	US-60-160-203-771	Sequence 771, App
74	144	5.3	728	70	US-60-160-203-838	Sequence 838, App
75	123	4.5	123	36	US-09-864-761-28381	Sequence 28381, A
76	123	4.5	123	47	US-10-182-993-20377	Sequence 20377, A
77	123	4.5	123	48	US-10-203-135-20831	Sequence 20831, A

78	123	4.5	123	48	US-10-203-136-21236	Sequence 21236, A
79	122	4.5	456	1	PCT-US01-00663-6904	Sequence 6904, Ap
80	122	4.5	456	2	PCT-US01-00663-6904	Sequence 6904, Ap
81	122	4.5	456	36	US-09-864-761-6056	Sequence 6056, Ap
82	122	4.5	456	47	US-10-182-993-6667	Sequence 6667, Ap
83	122	4.5	456	47	US-10-182-993-6056	Sequence 6056, Ap
84	122	4.5	456	48	US-10-203-134-7146	Sequence 7146, Ap
85	122	4.5	456	48	US-10-203-135-7929	Sequence 7929, Ap
86	122	4.5	456	48	US-10-203-136-7070	Sequence 7070, Ap
87	122	4.5	456	48	US-10-203-137-6904	Sequence 6904, Ap
88	122	4.5	456	48	US-10-203-139-6862	Sequence 6862, Ap
89	122	4.5	555	72	US-60-181-996-169	Sequence 169, App
90	115	4.2	588	71	US-60-177-571-733	Sequence 733, App
91	109	4.0	405	80	US-60-252-833-39848	Sequence 39848, A
92	103	3.8	385	71	US-60-177-571-962	Sequence 962, App
93	103	3.8	385	71	US-60-177-571-963	Sequence 963, App
94	101	3.7	101	1	PCT-US01-00663-20219	Sequence 20219, A
95	101	3.7	101	2	PCT-US01-00663-20219	Sequence 20219, A
96	101	3.7	101	36	US-09-864-761-22893	Sequence 22893, A
97	101	3.7	101	47	US-10-182-993-19584	Sequence 19584, A
98	101	3.7	101	47	US-10-182-993-16039	Sequence 16039, A
99	101	3.7	101	48	US-10-203-134-20153	Sequence 20153, A
100	101	3.7	101	48	US-10-203-135-19859	Sequence 19859, A

ALIGNMENTS

RESULT 1
US-09-774-490-1
; Sequence 1, Application US/09774490
; GENERAL INFORMATION:
; APPLICANT: Jin, Shengfang
; TITLE OF INVENTION: RESISTANCE SEQUENCES AND USES THEREOF
; FILE REFERENCE: 07334-138001
; CURRENT APPLICATION NUMBER: US/09/774,490
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: US 60/179,191
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows, Version 4.0
; SEQ ID NO 1
; LENGTH: 2709
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-774-490-1

Query Match 100.0%; Score 2709; DB 33; Length 2709;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2709; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AATCTTATTTATGATGTTAACAAGCTTAGTAATGATGCCAGGGGTGCG	60
DB	1	AATCTTATTTATGATGTTAACAAGCTTAGTAATGATGCCAGGGGTGCG	60
QY	61	ACCAAGCGGTGCGAGTAGTGGCTGCTGCTCCATTGTCAGCAGTATTT	120
DB	61	ACCAAGCGGTGCGAGTAGTGGCTGCTGCTCCATTGTCAGCAGTATTT	120
QY	121	CCAGTGTGTTGAATCTTCTGCGCCGACATPACAGAGAAGACTAAGCAGAAG	180
DB	121	CCAGTGTGTTGAATCTTCTGCGCCGACATPACAGAGAAGACTAAGCAGAAG	180
QY	181	GACCAAGCGGTGCGAGTAGTGGCTGCTGCTCCATTGTCAGCAGTATTT	240
DB	181	GACCAAGCGGTGCGAGTAGTGGCTGCTGCTCCATTGTCAGCAGTATTT	240
QY	241	ATTACTTACACCAAGAGCAACTATCAGATGGAGAGACATGTGCCAGGCTGAATT	300
DB	241	ATTACTTACACCAAGAGCAACTATCAGATGGAGAGACATGTGCCAGGCTGAATT	300
QY	301	ATCTCAAGAAATGTTGATCCACATGTGATCATTTCAATGGCTTGCCACAG	360
DB	301	ATCTCAAGAAATGTTGATCCACATGTGATCATTTCAATGGCTTGCCACAG	360

Db 301 ATCTTAAGAAAGTGGAAATCCAAATGATGATCACTTTCATGCTGGCCAAAG 360
 Qy 361 CTCGATTTATACCTTCTTTGGATGAGAAACGAGTAGCTGTATGTTGAGCAAA 420
 Db 361 CTCGATTTATACCTTCTTTGGATGAGAAACGAGTAGCTGTATGTTGAGCAAA 420
 Qy 421 GGATACATATTTTCATTCGACCTGGTTAAATCAAGATTTTCAAAAGATTTGTGGCC 480
 Db 421 GGATACATATTTTCATTCGACCTGGTTAAATCAAGATTTTCAAAAGATTTGTGGCC 480
 Qy 481 ACTATCTTACACAGAAAGATGAAATGCAAGTGGGTGAAAAAGACATCTTAAAGAAATG 540
 Db 481 ACTATCTTACACAGAAAGATGAAATGCAAGTGGGTGAAAAAGACATCTTAAAGAAATG 540
 Qy 541 TGTCTATTTTCATCAAGTACTTAAGCATATATCACTCACTTGTACGCTGTGAAAC 600
 Db 541 TGTCTATTTTCATCAAGTACTTAAGCATATATCACTCACTTGTACGCTGTGAAAC 600
 Qy 601 GGGGCTTTTCATCCAAATTTGCACTTCAATGAAATGGAATCATCTGAGAGCAATAT 660
 Db 601 GGGGCTTTTCATCCAAATTTGCACTTCAATGAAATGGAATCATCTGAGAGCAATAT 660
 Qy 661 TTTTAAAGCTGGAAGACTCACTTTTGAAGAAACGCGGTGGAGAGATCCATATGACCTTA 720
 Db 661 TTTTAAAGCTGGAAGACTCACTTTTGAAGAAACGCGGTGGAGAGATCCATATGACCTTA 720
 Qy 721 GCTGCTGACAGATCCCTTTTAAATGATGAGAAATTAATCTGGAAGCTGAGCTGATTT 780
 Db 721 GCTGCTGACAGATCCCTTTTAAATGATGAGAAATTAATCTGGAAGCTGAGCTGATTT 780
 Qy 781 TATGGGGGAGACTTTGCTATCTCCGAATCTTGGGACCAACCAACCAATCAGAGACA 840
 Db 781 TATGGGGGAGACTTTGCTATCTCCGAATCTTGGGACCAACCAACCAATCAGAGACA 840
 Qy 841 GAGAGATGATTTCCAGTGGCTCAATGATGCAAAAGTTCAATTAAGTCCCATCTCATCAGA 900
 Db 841 GAGAGATGATTTCCAGTGGCTCAATGATGCAAAAGTTCAATTAAGTCCCATCTCATCAGA 900
 Qy 901 GAGTACCAATCTGGAAGATGACAAAGTATATCTTTTCTCGGTGAAGAAATGCAATGATG 960
 Db 901 GAGTACCAATCTGGAAGATGACAAAGTATATCTTTTCTCGGTGAAGAAATGCAATGATG 960
 Qy 961 AGAACAATCTTGGAAAAGCTACTCAGCTGATGAAATAGGTCAGATATGCAAGATGACTTTGG 1020
 Db 961 AGAACAATCTTGGAAAAGCTACTCAGCTGATGAAATAGGTCAGATATGCAAGATGACTTTGG 1020
 Qy 1021 AGGGAACAGAAATCTGGTGAATTAATGCAACATTTCTCAAAAGCTCGTGTGATTTGGCTC 1080
 Db 1021 AGGGAACAGAAATCTGGTGAATTAATGCAACATTTCTCAAAAGCTCGTGTGATTTGGCTC 1080
 Qy 1081 AGTGCAGAGTCCAAATGGATGACATCTATTTGATGAACTGCAAGATGATGCTAT 1140
 Db 1081 AGTGCAGAGTCCAAATGGATGACATCTATTTGATGAACTGCAAGATGATGCTAT 1140
 Qy 1141 GAATTTAAAGATCTTAAATATCAAGTGTATGATGAGATTTTGAAGTCTTCAAGTAAGAT 1200
 Db 1141 GAATTTAAAGATCTTAAATATCAAGTGTATGATGAGATTTTGAAGTCTTCAAGTAAGAT 1200
 Qy 1201 TTTTCAAGGATGAGCGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
 Db 1201 TTTTCAAGGATGAGCGGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
 Qy 1261 TCCATATGCCCAAGAGATGAGCCCAATATGAGGAGGCTTTCATCAAGAGAGAGTCCC 1320
 Db 1261 TCCATATGCCCAAGAGATGAGCCCAATATGAGGAGGCTTTCATCAAGAGAGAGTCCC 1320
 Qy 1321 CTATCCAGGCGCAGGAATCTTGTCCAGCAAAATTTTGTGTTTGAATCTCAAGAAAGTA 1380
 Db 1321 CTATCCAGGCGCAGGAATCTTGTCCAGCAAAATTTTGTGTTTGAATCTCAAGAAAGTA 1380
 Qy 1381 CCTTCTGATGATGTTATTAACCTTTGCAAGAAAGTATCCAGGCAATGATCAATCAAGTGT 1440
 Db 1381 CCTTCTGATGATGTTATTAACCTTTGCAAGAAAGTATCCAGGCAATGATCAATCAAGTGT 1440

Qy 1441 TCTATGAACAATGCGCCCAATAGTATGATCAAAAAGATGTAATTAATTAATTAACAAAT 1500
 Db 1441 TCTATGAACAATGCGCCCAATAGTATGATCAAAAAGATGTAATTAATTAATTAACAAAT 1500
 Qy 1501 TGTGTGAGACCGAGTGTATGCAAGAAATGACAGTATGATGATGATGATGATGATGATGATG 1560
 Db 1501 TGTGTGAGACCGAGTGTATGCAAGAAATGACAGTATGATGATGATGATGATGATGATGATG 1560
 Qy 1561 TGTGTGAGACCGTCTTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
 Db 1561 TGTGTGAGACCGTCTTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
 Qy 1621 GGTTCGTGGAAGAAATGACAGTGTTCGGGAAACCGACTGTATTTTCAAGAAATGAGACT 1680
 Db 1621 GGTTCGTGGAAGAAATGACAGTGTTCGGGAAACCGACTGTATTTTCAAGAAATGAGACT 1680
 Qy 1681 TTCCATGAGCAGCAACAATATATATGTTTCAACGCGTGGGGTGGCCAGCTCCCTTT 1740
 Db 1681 TTCCATGAGCAGCAACAATATATATGTTTCAACGCGTGGGGTGGCCAGCTCCCTTT 1740
 Qy 1741 AACCGGTGTATATTTACGGGAAAGCTGTGAGTGTGCTGCGCCGAGACCTTTA 1800
 Db 1741 AACCGGTGTATATTTACGGGAAAGCTGTGAGTGTGCTGCGCCGAGACCTTTA 1800
 Qy 1801 CTGTGCTTGGGATGGTCTGATGTTCTGCTATTTTCCACTGCAAGAGACGCAAG 1860
 Db 1801 CTGTGCTTGGGATGGTCTGATGTTCTGCTATTTTCCACTGCAAGAGACGCAAG 1860
 Qy 1861 ACGACAGATATTAAGAAATGAGAACCACTGATCTCATCTGTTCAGACTTACACATGATTA 1920
 Db 1861 ACGACAGATATTAAGAAATGAGAACCACTGATCTCATCTGTTCAGACTTACACATGATTA 1920
 Qy 1921 TCAACATGCGCACAGCCCTGAAAGAGAAATCATCTATGCTGTAGAAATAGTAGCACTT 1980
 Db 1921 TCAACATGCGCACAGCCCTGAAAGAGAAATCATCTATGCTGTAGAAATAGTAGCACTT 1980
 Qy 1981 TTTGGAATGACGTCCGAAGTGCAGAGAGCCGTGCTATTTGGCAATTCAGAGGCGAAA 2040
 Db 1981 TTTGGAATGACGTCCGAAGTGCAGAGAGCCGTGCTATTTGGCAATTCAGAGGCGAAA 2040
 Qy 2041 TGAAGAGGAAAGAGATCAGATGATGATCATATCATCAGACAGATCAAGGCT 2100
 Db 2041 TGAAGAGGAAAGAGATCAGATGATGATCATATCATCAGACAGATCAAGGCT 2100
 Qy 2101 TCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160
 Db 2101 TCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2160
 Qy 2161 TGGGTTCATCAAACTCTTTTAAAGTAAACCTGGAAGTCAATGACAGAGCAATTTGGA 2220
 Db 2161 TGGGTTCATCAAACTCTTTTAAAGTAAACCTGGAAGTCAATGACAGAGCAATTTGGA 2220
 Qy 2221 AGAATCTTCTCATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2280
 Db 2221 AGAATCTTCTCATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2280
 Qy 2281 CATGACACTTACCGAAGAGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2340
 Db 2281 CATGACACTTACCGAAGAGTCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 2340
 Qy 2341 TCTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
 Db 2341 TCTCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2400
 Qy 2401 GCAAAGGCGCAGAGATACCCCAAGGAAACATTAATGAAAGCACTTACAAAGAAATTA 2460
 Db 2401 GCAAAGGCGCAGAGATACCCCAAGGAAACATTAATGAAAGCACTTACAAAGAAATTA 2460
 Qy 2461 GAAAGGTGAAACAGAGAGATCCCAAGATTTTGAAGGCAACCCAGAGTGTCTGAGCTGC 2520
 Db 2461 GAAAGGTGAAACAGAGAGATCCCAAGATTTTGAAGGCAACCCAGAGTGTCTGAGCTGC 2520

```
QY 2521 ATTACCTTAGAAGAACTCAAGAGTAGAAGAACTTCCCTAGACATAAATGGAAGAAACAA 2580
DB 2521 ATTACCTTAGAAGAACTCAAGAGTAGAAGAACTTCCCTAGACATAAATGGAAGAAACAA 2580
QY 2581 TGAATATATCAATGAACTTTTTCATGAGCTATATGAGAGTTTAAATGAGTGGGAATTC 2640
DB 2581 TGAATATATCAATGAACTTTTTCATGAGCTATATGAGAGTTTAAATGAGTGGGAATTC 2640
QY 2641 AGCTGAGTTCCACCAATTAATTAATCATGAGTAACCTTCTTAATAGCTTTTTC 2700
DB 2641 AGCTGAGTTCCACCAATTAATTAATCATGAGTAACCTTCTTAATAGCTTTTTC 2700
QY 2701 CCTAATACC 2709
DB 2701 CCTAATACC 2709

RESULT 2
US-10-247-671-83
; Sequence 83, Application US/10247671
; GENERAL INFORMATION:
; APPLICANT: Mikita, Thomas
; APPLICANT: Shiffman, Dov
; APPLICANT: Porter, Gordon, J.
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
; FILE REFERENCE: PA-0050 US
; CURRENT APPLICATION NUMBER: US/10/247,671
; PRIOR FILING DATE: 2002-09-18
; PRIOR APPLICATION NUMBER: 60/323,784
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PERL Program
; SEQ ID NO 83
; LENGTH: 2848
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: 1930967CB1
US-10-247-671-83

Query Match 92.9%; Score 2517; DB 48; Length 2848;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2517; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 CTGCGACATGGGCTGGTTAATCTAGATGTCGTCTCTTTCTGGGGAGTATTAATTACACAGC 252
DB 9 CTGCGACATGGGCTGGTTAATCTAGATGTCGTCTCTTTCTGGGGAGTATTAATTACACAGC 68
QY 251 AAGAGCAACTATCAGAAATGGAAGAAACAATGTGCCAAGGCTGGAATTAATCTCAACAAGA 312
DB 69 AAGAGCAACTATCAGAAATGGAAGAAACAATGTGCCAAGGCTGGAATTAATCTCAACAAGA 128
QY 313 AATGTTGGAATCCACAATGTGATCACTTTCAATGSCCTTGCCACAAGCTCCAGTTATCA 372
DB 129 AATGTTGGAATCCACAATGTGATCACTTTCAATGSCCTTGCCACAAGCTCCAGTTATCA 188
QY 373 TACCTTCCCTTTGAGAGAGAAAGGAGTGGCTGTATGTTGAGCAAAAGATCACATATT 432
DB 189 TACCTTCCCTTTGAGAGAGAAAGGAGTGGCTGTATGTTGAGCAAAAGATCACATATT 248
QY 433 TTCAATTCGACCTGTTAATATCAAGGATTTTCAAAAGATGTGTGCGCAGTATCTTACAC 492
DB 249 TTCAATTCGACCTGTTAATATCAAGGATTTTCAAAAGATGTGTGCGCAGTATCTTACAC 308
QY 493 CAGAAGAGATGATGCAAGTGGGCTGGAAGAAAGACATCTGGAAGAAATGTGCTAATTTCA 552
DB 309 CAGAAGAGATGATGCAAGTGGGCTGGAAGAAAGACATCTGGAAGAAATGTGCTAATTTCA 368
QY 553 CAAAGTACTTAAGGATTAATATCAAGTCACTTTGTACGCTGTGGAACGGGGGCTTTTCA 612
DB 369 CAAAGTACTTAAGGATTAATATCAAGTCACTTTGTACGCTGTGGAACGGGGGCTTTTCA 428
```

```
QY 613 TCCAAATTCGACCTTACATTTGAATGAGCATATCTTGAGAGCAATATTTTAACTGGA 672
DB 429 TCCAAATTCGACCTTACATTTGAATGAGCATATCTTGAGAGCAATATTTTAACTGGA 488
QY 673 GAATCTACATTTTGAAGAACGGCGGTGGAGAGTCAATATGACCTTAAGCTGCTACAGC 732
DB 489 GAATCTACATTTTGAAGAACGGCGGTGGAGAGTCAATATGACCTTAAGCTGCTACAGC 548
QY 733 ATCCCTTTTAATAGATGAGAAATTAATGATGAGCTGGAAGTCAAGCTGATTTTATGGGCGAGA 792
DB 549 ATCCCTTTTAATAGATGAGAAATTAATGATGAGCTGGAAGTCAAGCTGATTTTATGGGCGAGA 608
QY 793 CTTTGCTATCTTCCGAACTTTGGGACACACCACCAATCAGAGCAGAGCATGATTC 852
DB 609 CTTTGCTATCTTCCGAACTTTGGGACACACCACCAATCAGAGCAGAGCATGATTC 668
QY 853 CAGGTGGCTCAATGATCCAAAGTTTATTAAGTCCCACTCATCTGAGAGTGAACATTC 912
DB 669 CAGGTGGCTCAATGATCCAAAGTTTATTAAGTCCCACTCATCTGAGAGTGAACATTC 728
QY 913 TGAAGATGCAAGATATCTTTTCTCCGTAAGAAATGCAATAGATGAGAAACACTGTGG 972
DB 729 TGAAGATGCAAGATATCTTTTCTCCGTAAGAAATGCAATAGATGAGAAACACTGTGG 788
QY 973 AAAAGCTACTCAGCTAGAAATAGTCAATATGCAAGATGACTTTGGAGGGCAGAGAAG 1032
DB 789 AAAAGCTACTCAGCTAGAAATAGTCAATATGCAAGATGACTTTGGAGGGCAGAGAAG 848
QY 1033 TTTGCTGAATTAATGAGCAACATTTCTCAAGGCTGTCTGATTTGCTCAAGTCCAGGTC 1092
DB 849 TTTGCTGAATTAATGAGCAACATTTCTCAAGGCTGTCTGATTTGCTCAAGTCCAGGTC 908
QY 1093 AAATGCAATGCACTCATTTTGTATGAGTCAAGTCAAGATGATATTCCTAATGAATTAAGA 1152
DB 909 AAATGCAATGCACTCATTTTGTATGAGTCAAGTCAAGATGATATTCCTAATGAATTAAGA 968
QY 1153 TCTTAAAAATCCAGTTGTATATGAGTGTATACGATTCAGATTAACATTTTCAAGGATC 1212
DB 969 TCTTAAAAATCCAGTTGTATATGAGTGTATACGATTCAGATTAACATTTTCAAGGATC 1028
QY 1213 AGCGGTGTATGTATAGATGAGTATGAGTGAAGAGGATGTTCTTGATCCATATAGCCCA 1272
DB 1029 AGCGGTGTATGTATAGATGAGTATGAGTGAAGAGGATGTTCTTGATCCATATAGCCCA 1088
QY 1273 CAGGATGAGCCCAACTATCAATGAGTCCCTTATCAAGAAAGATCCCTATCCAGGCT 1332
DB 1089 CAGGATGAGCCCAACTATCAATGAGTCCCTTATCAAGAAAGATCCCTATCCAGGCT 1148
QY 1333 AGGAATTTGTCAGCAAAACATTTGTTGACTCTCAAAAGACCTTCTGATGA 1392
DB 1149 AGGAATTTGTCAGCAAAACATTTGTTGACTCTCAAAAGACCTTCTGATGA 1208
QY 1393 TGTATTAACCTTTGCAAGAGTCAATCCAGCATGATCAATCAAGTGTTCCTATGAAACA 1452
DB 1209 TGTATTAACCTTTGCAAGAGTCAATCCAGCATGATCAATCAAGTGTTCCTATGAAACA 1268
QY 1453 TCGCCCAATAGATCAAAAGGATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1512
DB 1269 TCGCCCAATAGATCAAAAGGATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1328
QY 1513 AGTGAATCAGAGATGAGACATGATGATGTTATGCGAAACAGATGTTGGAGCCGT 1572
DB 1329 AGTGAATCAGAGATGAGACATGATGATGTTATGCGAAACAGATGTTGGAGCCGT 1388
QY 1573 TCTTAAGTATGTTCAATTTCTTAAGAGACCTTGTATGATTTTAAGAGAGTTCGTGGA 1632
DB 1389 TCTTAAGTATGTTCAATTTCTTAAGAGACCTTGTATGATTTTAAGAGAGTTCGTGGA 1448
QY 1633 AGAATGACAGTTTGGGAAACGAGTCTATTTAGCAATGAGGCTTCCACTAAGCA 1692
DB 1449 AGAATGACAGTTTGGGAAACGAGTCTATTTAGCAATGAGGCTTCCACTAAGCA 1508
```

OY	1693	GCAACAACTATATATTTGGTTCAAGGAGCTGGGGGTGGCCAGCTCCCTTTAAACCCGGTGGGA	1755
Db	1509	GCAACAACTATATATTTGGTTCAAGGAGCTGGGGGTGGCCAGCTCCCTTTAAACCCGGTGGGA	1566
OY	1753	TATTTACGGGAAAGCGTGTGCTGAAGTGTTCCTCGGCCGAGACCTTACTGTGCTTGGGA	1812
Db	1569	TATTTACGGGAAAGCGTGTGCTGAAGTGTTCCTCGGCCGAGACCTTACTGTGCTTGGGA	1628
OY	1873	AAGAAATGGAGACCCACTGACTCACTGTTCAAGCTTAAACCACTGATTAATCAACATAGGCA	1932
Db	1669	AAGAAATGGAGACCCACTGACTCACTGTTCAAGCTTAAACCACTGATTAATCAACATAGGCA	1748
OY	1933	CAGCCCTGAAGAGAGAAATCATCTATGTGTGAGAAATAGTAGCACATTTTGGAAATGACG	1992
Db	1749	CAGCCCTGAAGAGAGAAATCATCTATGTGTGAGAAATAGTAGCACATTTTGGAAATGACG	1806
OY	1993	TCCGAAGTCGCAAGAGAGCGCTGTGTTATTTGGCAATTTCCAGAGGCGAAATGAAAGCGGAA	2055
Db	1809	TCCGAAGTCGCAAGAGAGCGCTGTGTTATTTGGCAATTTCCAGAGGCGAAATGAAAGCGGAA	1866
OY	2053	AGAAGAGATCAGAGTGGATGTGATCATATATCAAGACAGATCAAGGCTCTTGCTACTGTAG	2112
Db	1869	AGAAGAGATCAGAGTGGATGTGATCATATATCAAGACAGATCAAGGCTCTTGCTACTGTAG	1922
OY	2113	TCTACAAAGAGAGATTCAAGCAATTACGCTTGCAATGCGGTGGAAATGAGGTTTATACA	2172
Db	1939	TCTACAAAGAGAGATTCAAGCAATTACGCTTGCAATGCGGTGGAAATGAGGTTTATACA	1988
OY	2173	AACTCTTCTTAAGATTAACCTTGGAAGTCAATTGACAACAAGCATTTGGAGAACTTCTTCA	2232
Db	1989	AACTCTTCTTAAGATTAACCTTGGAAGTCAATTGACAACAAGCATTTGGAGAACTTCTTCA	2048
OY	2233	TAAAGATGATGATGGAGATGGCTCTAAGAACCAAAAGAAATGTCTCAATACATGACACCTAG	2292
Db	2049	TAAAGATGATGATGGAGATGGCTCTAAGAACCAAAAGAAATGTCTCAATACATGACACCTAG	2108
OY	2293	CCAGAGGTCGTGTACAGAGACTTCATGCACTCATCAACCCCAATCTTCAACAGAT	2352
Db	2109	CCAGAGGTCGTGTACAGAGACTTCATGCACTCATCAACCCCAATCTTCAACAGAT	2168
OY	2353	GGATGAGTCTGTGAAACAAGTTTGGAAAAAGGACCGAAAAACAACGTCCGCAAAAGGCCACG	2412
Db	2169	GGATGAGTCTGTGAAACAAGTTTGGAAAAAGGACCGAAAAACAACGTCCGCAAAAGGCCACG	2228
OY	2413	ACATACCCCAAGGGAACAGTTAACAAATGGAAGCACTTACAGAAATTAAGAAAGTATGAAA	2472
Db	2229	ACATACCCCAAGGGAACAGTTAACAAATGGAAGCACTTACAGAAATTAAGAAAGTATGAAA	2288
OY	2473	CAGAGAGACCCACGAATTTGAGAGGGACCCAGAGATGTCTGAGCTGCATTACTCTTGA	2532
Db	2289	CAGAGAGACCCACGAATTTGAGAGGGACCCAGAGATGTCTGAGCTGCATTACTCTTGA	2348
OY	2533	AACCCCAAAACAAGTAAAGCTTGCTAGCAATACTGGAAAAACAATGCAATTAACAT	2592
Db	2349	AACCCCAAAACAAGTAAAGCTTGCTAGCAATACTGGAAAAACAATGCAATTAACAT	2408
OY	2593	GAACTTTTTCATGACATTAATGTGATGTTTAAACAATGATGGAAATTCAGCTGATTTCA	2652
Db	2409	GAACTTTTTCATGACATTAATGTGATGTTTAAACAATGATGGAAATTCAGCTGATTTCA	2466
OY	2653	CCAATTAATAATTAATCAATGAGTAACCTTTCCTAATAGGCTTTTTCCTAATACC	2709
Db	2469	CCAATTAATAATTAATCAATGAGTAACCTTTCCTAATAGGCTTTTTCCTAATACC	2525

```

1  APPLICANT: Morris, Macdonald
2  APPLICANT: Lal, Preeti
3  APPLICANT: Deep Dinh
4  TITLE OR INVENTION: Method for the Identification of Sequence Polymorphisms Using
5  TITLE OR INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide Polymorp
6  FILE REFERENCE: GX-0006 P
7  CURRENT APPLICATION NUMBER: US/60/172,373
8  CURRENT FILING DATE: 1999-12-16
9  NUMBER OF SEQ ID NOS: 25, 172
10 SOFTWARE: PERL Program
11 SEQ ID NO 8935
12 LENGTH: 2648
13 TYPE: DNA
14 ORGANISM: Homo sapiens
15 FEATURE:
16 NAME/KEY: misc_feature
17 OTHER INFORMATION: Incyte ID No: 337522.1
18 US-60-172-373-8935

```

Query Match	92.9%;	Score 2517;	DB 71;	Length 2848;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 2517;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Oy	193	CTGAGCATGGGCTGTTAACTAGGATGTCTGTCTTTTCTGGGGAGTTTACTTACAC	252
Db	9	CTGCAGCATGGGCTGTTAACTAGGATGTCTGTCTTTTCTGGGGAGTTTACTTACAC	68
Oy	253	AAGAGCAACTATACGAATGGGAGAACAAATGTGCGAAAGGCTGAATTTATCTTACAAAG	312
Db	69	AAGAGCAAACTATACGAATGGGAGAACAAATGTGCGAAAGGCTGAATTTATCTTACAAAG	128
Oy	313	AATGTGGAAATCCAAATGTATGATCACTTTCAAATGGCTTTGGGCAACGCTCCAGTTATCA	372
Db	129	AATGTGGAAATCCAAATGTATGATCACTTTCAAATGGCTTTGGGCAACGCTCCAGTTATCA	188
Oy	373	TACCTTCCTTTGGATGAGAACGGAGTAGGCTGTATGTGTGAGCAAAAGATCAACATAT	432
Db	189	TACCTTCCTTTGGATGAGAACGGAGTAGGCTGTATGTGTGAGCAAAAGATCAACATAT	248
Oy	433	TTCAATGACCTGGTTAATATCAGAGATTTTCAAAAAGATGTGTGGCCAGTATCTTACAC	492
Db	249	TTCAATTCACCTGGTTAATATCAGAGATTTTCAAAAAGATGTGTGGCCAGTATCTTACAC	308
Oy	493	CAGAGAGATGAATGCAATGGGCGCTGGAAAAAGACATCTCGAAGAAATGTGCTAATTTCAAT	552
Db	309	CAGAGAGATGAATGCAATGGGCGCTGGAAAAAGACATCTCGAAGAAATGTGCTAATTTCAAT	368
Oy	553	CAAGGTACTTAAGGCATTAATACAGACTCATTTGTAGCGCTGTGNAACGGGGGCTTTTCA	612
Db	369	CAAGGTACTTAAGGCATTAATACAGACTCATTTGTAGCGCTGTGNAACGGGGGCTTTTCA	428
Oy	613	TCCAAATTTGCACTATTAATGAAATTTGAGCATATCTCTGAGGACATATTTTAAAGCTGGA	672
Db	429	TCCAAATTTGCACTATTAATGAAATTTGAGCATATCTCTGAGGACATATTTTAAAGCTGGA	488
Oy	673	GAACTCAATTTTGA AAAACGGCGGTGGGAAAGTGCATATGACCTTAAGCTCTGTGACAG	732
Db	489	GAACTCAATTTTGA AAAACGGCGGTGGGAAAGTGCATATGACCTTAAGCTCTGTGACAG	548
Oy	733	ATCCCTTTTAATAGATGAGAAATTAATCTGTGAACTGACGCTGATTTTATGGGGCGAGA	792
Db	549	ATCCCTTTTAATAGATGAGAAATTAATCTGTGAACTGACGCTGATTTTATGGGGCGAGA	608
Oy	793	CTTTGGCTATCTTCGGAATCTTTGGGACACACCAACCATTCAGGAGCAGAGCAGATGATTC	852
Db	609	CTTTGGCTATCTTCGGAATCTTTGGGACACACCAACCATTCAGGAGCAGAGCAGATGATTC	668
Oy	853	CAGGTGGCTCATGATCCAAAGTTTCATTAATGTGCCACCTCATCTCAGAGAGTGAACAATCC	912
Db	669	CAGGTGGCTCATGATCCAAAGTTTCATTAATGTGCCACCTCATCTCAGAGAGTGAACAATCC	728
Oy	913	TGAAGATGACAAAGTATCTTTTCTTCGTTGAAAATGCAATATGATGAGAACACTCTGG	972

Db	729	TGAAATGACAAAGTAATCTTTTCTTCGCTGAATAATGAAATGAGAACTCTCG	788
Qy	973	AAAAGCTACTCAGCGTAGAATAGCTCAGATATGCAAGATGACTTTGGAGGGCACAGAG	10320
Db	789	AAAACCTACTCAGCGTAGAATAGCTCAGATATGCAAGATGACTTTGGAGGGCACAGAG	848
Qy	1033	TCTGTGTAATAAATGAGCAACTTCCTCAAACTGCTGATTTGCTCAGTGGCCAGGCTC	10920
Db	849	TCGTGTGAATAAATGAGCAACTTCCTCAAACTGCTGATTTGCTCAGTGGCCAGGCTC	908
Qy	1093	AAATGGCAATTGCACTCATTTTGTATGAACTGCAAGATGATTTCTTAATGAACCTTAAAGA	11520
Db	909	AAATGGCAATTGCACTCATTTTGTATGAACTGCAAGATGATTTCTTAATGAACCTTAAAGA	968
Qy	1153	TCCTTAAAAATCCAGTTGTATATGAGAGTGTATTCGACTTCAGTAACATTTTCAAGGATC	12120
Db	969	TCCTTAAAAATCCAGTTGTATATGAGAGTGTATTCGACTTCAGTAACATTTTCAAGGATC	10280
Qy	1213	AGCGGTGTATATGATATGATAGATAGATGTGAGAGGGGTTCCTTGGTCCATATNGCCA	12720
Db	1029	AGCGGTGTATATGATATGATAGATAGATGTGAGAGGGGTTCCTTGGTCCATATNGCCA	10880
Qy	1273	CAGGATGSAACCCCACTATCAATGGGTGCTTATCAAGSAGAGTCCCTATCCACGGCC	13320
Db	1089	CAGGATGSAACCCCACTATCAATGGGTGCTTATCAAGSAGAGTCCCTATCCACGGCC	11480
Qy	1333	AGGAATTGTCCACAGAAACATTTGGTGTTTGACTCTACAAAGACCTTCTGATGA	13920
Db	1149	AGGAATTGTCCACAGAAACATTTGGTGTTTGACTCTACAAAGACCTTCTGATGA	12080
Qy	1393	TGTTATTAACCTTTGCAAGAAAGTATCCAGCCATGACATCCAGTGTTCCTATGAAACA	14520
Db	1209	TGTTATTAACCTTTGCAAGAAAGTATCCAGCCATGACATCCAGTGTTCCTATGAAACA	12680
Qy	1453	TCGCCAATATGATCAAAACCGSAGTAAATTAATCAATTTACAAATTTGTCGTAGACCG	15120
Db	1269	TCGCCAATATGATCAAAACCGSAGTAAATTAATCAATTTACAAATTTGTCGTAGACCG	13280
Qy	1513	AGTGATGCAGAAAGTGAACAGTATGATGTATGTATTCGGAACAGATGTTGGGACCGT	15720
Db	1329	AGTGATGCAGAAAGTGAACAGTATGATGTATGTATTCGGAACAGATGTTGGGACCGT	13880
Qy	1573	TCTTAATATGTTTCAATTTCTTANAGSAACTTGTATATTAAGAAGGTTCTGCTGGA	16320
Db	1389	TCTTAATATGTTTCAATTTCTTANAGSAACTTGTATATTAAGAAGGTTCTGCTGGA	14480
Qy	1633	AGAAATGACAGTTTTCGGSAAACGCACTGCTAATTTCAAGCAATGAGCTTCCACTAAGCA	16920
Db	1449	AGAAATGACAGTTTTCGGSAAACGCACTGCTAATTTCAAGCAATGAGCTTCCACTAAGCA	15080
Qy	1693	GCAACAACTATATATTTGGTTCAACGGCTGGGGGTTCGCCAGCTCCCTTTACACGGTGTGA	17520
Db	1509	GCAACAACTATATATTTGGTTCAACGGCTGGGGGTTCGCCAGCTCCCTTTACACGGTGTGA	15680
Qy	1753	TATTTACGGGAAAGGTGTGCTGATGTGTGCTCGCCCGGAAACCTTAATCTGTCTGGGA	18120
Db	1569	TATTTACGGGAAAGGTGTGCTGATGTGTGCTCGCCCGGAAACCTTAATCTGTCTGGGA	16280
Qy	1813	TGTTTCTGCAATGTTCTCGCTAATTTTCCCACTGCAAAAGAGCGCAAGAGCAAGAATAT	18720
Db	1629	TGTTTCTGCAATGTTCTCGCTAATTTTCCCACTGCAAAAGAGCGCAAGAGCAAGAATAT	16880
Qy	1873	AAGAAATGAGAACCCACTGACTCATGTGTTCAGACTTAACCAATGATATATCAACATGGCA	19320
Db	1689	AAGAAATGAGAACCCACTGACTCATGTGTTCAGACTTAACCAATGATATATCAACATGGCA	17480
Qy	1933	CAGCCCTGAAAGAGATCATCTATGCTGTATGAGAAATGTACACATTTTGGAAATGAG	19920
Db	1749	CAGCCCTGAAAGAGATCATCTATGCTGTATGAGAAATGTACACATTTTGGAAATGAG	18080
Qy	1993	TCCGAAGTCGAGAGAGCGCTGTCTATTTGGCAATTTCCAGAGGGGAAATGAAAGCGAAA	20520
Db	1809	TCCGAAGTCGAGAGAGCGCTGTCTATTTGGCAATTTCCAGAGGGGAAATGAAAGCGAAA	18680

QY	2053	AGAAAGATCA	GAGTGA	TGATTC	ATCATCAT	CAGGACA	CAATCA	AGGCGCTT	GTCTG	CTAG	2113		
Db	1869	AGAAAGATCA	GAGTGA	TGATTC	ATCATCAT	CAGGACA	CAATCA	AGGCGCTT	GTCTG	CTAG	1922		
QY	2113	TCTCA	CAACAG	AGATTC	CAGGCA	ATTAC	GGCAAT	TTCCT	CTCCAT	CGGTGG	AACTGGGTT	2172	
Db	1929	TCTCA	CAACAG	AGATTC	CAGGCA	ATTAC	GGCAAT	TTCCT	CTCCAT	CGGTGG	AACTGGGTT	1988	
QY	2173	AACTCTT	TTAAGT	TAACCT	CTGGA	AGTCAT	TGACA	CAGAGCA	ATTGG	AAAGAACTT	CTTCA	2233	
Db	1989	AACTCTT	TTAAGT	TAACCT	CTGGA	AGTCAT	TGACA	CAGAGCA	ATTGG	AAAGAACTT	CTTCA	2044	
QY	2233	TAAAGAT	GATGAT	GGAGAT	GGCTCT	TAAGAC	CCAAAGAA	TGTCCA	TATGAC	TAGCA	CACTTG	2299	
Db	2049	TAAAGAT	GATGAT	GGAGAT	GGCTCT	TAAGAC	CCAAAGAA	TGTCCA	TATGAC	TAGCA	CACTTG	2108	
QY	2293	CCAGAGG	TCGTG	TACAG	AGACTT	CATGAC	TCTAT	CAACCA	CCCCCA	TTCTCA	CACAT	2355	
Db	2109	CCAGAGG	TCGTG	TACAG	AGACTT	CATGAC	TCTAT	CAACCA	CCCCCA	TTCTCA	CACAT	2168	
QY	2353	GGATGAG	TTCGT	GTAA	CAAGTT	TGGAAA	AAAGGGA	CCGAAA	CAAGTGGG	CAAAAGG	CCAGG	2412	
Db	2169	GGATGAG	TTCGT	GTAA	CAAGTT	TGGAAA	AAAGGGA	CCGAAA	CAAGTGGG	CAAAAGG	CCAGG	2222	
QY	2413	ACATAT	CCCCC	AGAG	CACTTA	CAAAATGGA	AGAC	CTTCA	CAAGAAA	ATTA	GAAAAGT	2472	
Db	2229	ACATAT	CCCCC	AGAG	CACTTA	CAAAATGGA	AGAC	CTTCA	CAAGAAA	ATTA	GAAAAGT	2288	
QY	2473	CAGAG	AGAC	CCAC	GAATTT	TGAG	AGGGCA	CCCGAG	AGTGTCT	GAGTCA	TTAACCT	CTTGA	2533
Db	2289	CAGAG	AGAC	CCAC	GAATTT	TGAG	AGGGCA	CCCGAG	AGTGTCT	GAGTCA	TTAACCT	CTTGA	2348
QY	2533	AACTCA	AAACAG	TAGAA	ACTTGG	CTCGA	CAATAT	ACGTGA	AAAAA	CAATAT	GCATTA	CAT	2592
Db	2349	AACTCA	AAACAG	TAGAA	ACTTGG	CTCGA	CAATAT	ACGTGA	AAAAA	CAATAT	GCATTA	CAT	2408
QY	2593	GAAC	TTTTTT	CATG	GCATTA	TATG	GGATGTT	TACAT	GGTGG	AAATTC	CAGTGA	AGTTCA	2652
Db	2409	GAAC	TTTTTT	CATG	GCATTA	TATG	GGATGTT	TACAT	GGTGG	AAATTC	CAGTGA	AGTTCA	2468
QY	2653	CCAAT	TTATAA	TTAAAT	CCATGA	BAATCTT	CCATAT	AGGCTTTT	TTTCC	TAATAC	2709		
Db	2469	CCAAT	TTATAA	TTAAAT	CCATGA	BAATCTT	CCATAT	AGGCTTTT	TTTCC	TAATAC	2525		

RESULT 4
US-60-323-784-83

```

GENERAL_INFORMATION:
APPLICANT: Mikita, Thomas
APPLICANT: Shiffman, Dov
APPLICANT: Porter, Gordon, J.
APPLICANT: Kaseer, Matthew R.
TITLE OF INVENTION: GENES EXPRESSED IN TREATED FOAM CELLS
FILE REFERENCE: PA-0050 P
CURRENT APPLICATION NUMBER: US/60/323,784
CURRENT FILING DATE: 2001-09-19
NUMBER OF SEQ. ID NOS.: 186
SOFTWARE: PERL Program
SEQ ID NO 83
LENGTH: 2848
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 1930967CBI
US-60-323-784-83

```

Query Match	92.9%	Score 2517	DB 87	Length 2846
Best Local Similarity	100.0%	Pred. No. 0		
Matches 2517; Conservative	0	Mismatches	0	Indels 0; Gaps 0;

QY 193 CTGCGCATGGGCTGGTACTAGATGTTGTCTTTCTTGGGAGATTACTTAACAC 252
Db 9 CTGCGCATGGGCTGGTACTAGATGTTGTCTTTCTTGGGAGATTACTTAACAC 68
QY 253 AAGACAACATCTATGAAATGGGAGAACAAATGGCCAGGCTGAATTAATCTTACAAAGA 312
Db 69 AAGACAACATCTATGAAATGGGAGAACAAATGGCCAGGCTGAATTAATCTTACAAAGA 128
QY 313 AATGTTGGAATCCAAATGTTGATCATCTTCAATGGCTGGCCAAACAGCTCCAGTTATCA 372
Db 129 AATGTTGGAATCCAAATGTTGATCATCTTCAATGGCTGGCCAAACAGCTCCAGTTATCA 188
QY 373 TACCTTCTTTTGGATGAGGAGCGAGTACCTGATGTTGGAGCAAAAGATCACATATT 432
Db 189 TACCTTCTTTTGGATGAGGAGCGAGTACCTGATGTTGGAGCAAAAGATCACATATT 248
QY 433 TTCAATTCGACCTGTTAATATCAAGATTTTCAAAAGATGTTGTTGAGATTAATCTTAAC 492
Db 249 TTCAATTCGACCTGTTAATATCAAGATTTTCAAAAGATGTTGTTGAGATTAATCTTAAC 308
QY 493 CAGAGAGATGATCAAGTGGGCTGGAAGACATCCCTGAAGAGATGCTTAATTTCAAT 552
Db 309 CAGAGAGATGATCAAGTGGGCTGGAAGACATCCCTGAAGAGATGCTTAATTTCAAT 368
QY 553 CAGGTTACTTAAAGCATATATCACTCACTTGTACCGCTGTGGAACGGGGCTTTTCA 612
Db 369 CAGGTTACTTAAAGCATATATCACTCACTTGTACCGCTGTGGAACGGGGCTTTTCA 428
QY 613 TTCAATTCGACCTGTTAATATCAAGATTTTCAAAAGATGTTGTTGAGATTAATCTTAAC 672
Db 429 TTCAATTCGACCTGTTAATATCAAGATTTTCAAAAGATGTTGTTGAGATTAATCTTAAC 488
QY 673 GAACTCACATTTTGAAGAAACGGGCTGGAGAGAGTCCATATGACCTTAAGCTGTACAGC 732
Db 489 GAACTCACATTTTGAAGAAACGGGCTGGAGAGAGTCCATATGACCTTAAGCTGTACAGC 548
QY 733 ATCCCTTTTAAATAGAGAGATTAATCTTGTGAACCTGACGCTGATTTTATGGGGCAGA 792
Db 549 ATCCCTTTTAAATAGAGAGATTAATCTTGTGAACCTGACGCTGATTTTATGGGGCAGA 608
QY 793 CTGTTGCTATCTTCCGAACCTTGGGACCAACCCATTCAGAGAGAGAGAGATTTTC 852
Db 609 CTGTTGCTATCTTCCGAACCTTGGGACCAACCCATTCAGAGAGAGAGATTTTC 668
QY 853 CAGGTGCTCATATGATCCAAAGTTTATGTCCTCACTCATCTCAGAGAGATCAATCC 912
Db 669 CAGGTGCTCATATGATCCAAAGTTTATGTCCTCACTCATCTCAGAGAGATCAATCC 728
QY 913 TGAAGATGACAAAGTATATCTTTTCTTCCGTGAAGATGATGATGAGAACACTCTGG 972
Db 729 TGAAGATGACAAAGTATATCTTTTCTTCCGTGAAGATGATGATGAGAACACTCTGG 788
QY 973 AAAAGCTACTCAGCGTATGATATGTCAGATATGCAAGATGACTTTTGAAGGAGACAGAG 1032
Db 789 AAAAGCTACTCAGCGTATGATATGTCAGATATGCAAGATGACTTTTGAAGGAGACAGAG 848
QY 1033 TCTGTGATTAATGAGCAACATCTCTCAAGCTGTCTGATTTCTCAGTCCAGGCTCC 1092
Db 849 TCTGTGATTAATGAGCAACATCTCTCAAGCTGTCTGATTTCTCAGTCCAGGCTCC 908
QY 1093 AAATGCGATGACACTCATTTTGTATGAACTGAGAGATGATTTCTTAATGAATTTTAAAGA 1152
Db 909 AAATGCGATGACACTCATTTTGTATGAACTGAGAGATGATTTCTTAATGAATTTTAAAGA 968
QY 1153 TCTTAATAAATCCAGTTGATATGAGAGTGTATGCACTTCCAGTAACTTTTCAAGGATC 1212
Db 969 TCTTAATAAATCCAGTTGATATGAGAGTGTATGCACTTCCAGTAACTTTTCAAGGATC 1028
QY 1213 AGCCGTGTATGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1272
Db 1029 AGCCGTGTATGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1088
QY 1273 CAGGATGAGACCAACTATCAATGAGGCTGCTTATCAAGAGAGATCCCTTATCCAGGCT 1332

Db 1089 CAGGATGAGACCAACTATCAATGAGGCTGCTTATCAAGAGAGATCCCTTATCCAGGCT 1148
QY 1333 AGGAATGTTGCTCCAGCAAAACATTTGGTGTGTTGACTCTCAAGAGAGCTTCCGATGA 1392
Db 1149 AGGAATGTTGCTCCAGCAAAACATTTGGTGTGTTGACTCTCAAGAGAGCTTCCGATGA 1208
QY 1393 TGTATTAACCTTTTCAAGAGATCATCCAGCATGATCAATCCAGTGTTCCTATGAACAA 1452
Db 1209 TGTATTAACCTTTTCAAGAGATCATCCAGCATGATCAATCCAGTGTTCCTATGAACAA 1268
QY 1453 TCCGCCAATGATGATCAAAAGGATGTAATTTATCAATTTACAAATTTGCTGAGACCG 1512
Db 1269 TCCGCCAATGATGATCAAAAGGATGTAATTTATCAATTTACAAATTTGCTGAGACCG 1328
QY 1513 AGTGAATGACAAAGATGAGACAGTATGATGATGATGATGATGATGATGATGATGATGAT 1572
Db 1329 AGTGAATGACAAAGATGAGACAGTATGATGATGATGATGATGATGATGATGATGATGAT 1388
QY 1573 TCTTAAGTATGTTCAATCTTCAAGAGACTTGTATGATTTAGAGAGTGTCTGCTGA 1632
Db 1389 TCTTAAGTATGTTCAATCTTCAAGAGACTTGTATGATTTAGAGAGTGTCTGCTGA 1448
QY 1633 AGAATGACAGTGTGTTGGGAAACCGACTGCTATTTACGAATGAGCTTTCACTAAGCA 1692
Db 1449 AGAATGACAGTGTGTTGGGAAACCGACTGCTATTTACGAATGAGCTTTCACTAAGCA 1508
QY 1693 GCAACAACTATATATGTTGTTCAAGGCTGGGGTGGCCAGCTCCCTTACACGGGTGGA 1752
Db 1509 GCAACAACTATATATGTTGTTCAAGGCTGGGGTGGCCAGCTCCCTTACACGGGTGGA 1568
QY 1753 TATTTAAGGAGAAACGTTGCTGAGTGTGCTGCGCCAGAGACCTTACTGTCTGGGA 1812
Db 1569 TATTTAAGGAGAAACGTTGCTGAGTGTGCTGCGCCAGAGACCTTACTGTCTGGGA 1628
QY 1813 TGTGTTGCTATGTTCTGCTATTTTCCACTGCAAGAGACGCAACAGACGACATAT 1872
Db 1629 TGTGTTGCTATGTTCTGCTATTTTCCACTGCAAGAGACGCAACAGACGACATAT 1688
QY 1873 AAGAAATGAGAGACCACTGACTCACTGTTGCACTTACCATGATATATCAACATGAGCA 1932
Db 1689 AAGAAATGAGAGACCACTGACTCACTGTTGCACTTACCATGATATATCAACATGAGCA 1748
QY 1933 CAGCCTGAGAGAGAGATTCATCTATGTTGATGAGAAATGATGACATTTTGGATGAG 1992
Db 1749 CAGCCTGAGAGAGAGATTCATCTATGTTGATGAGAAATGATGACATTTTGGATGAG 1808
QY 1993 TCCGAAGTCCGAGAGAGCGCTGTCTATTTGGCAATTCGAGGCGAAATGAGAGCAAA 2052
Db 1809 TCCGAAGTCCGAGAGAGCGCTGTCTATTTGGCAATTCGAGGCGAAATGAGAGCAAA 1868
QY 2053 AGAAGATGACAGATGATGATATCATATCATGAGACAGATCAAGGCTTCTGCTACGATG 2112
Db 1869 AGAAGATGACAGATGATGATATCATATCATGAGACAGATCAAGGCTTCTGCTACGATG 1928
QY 2113 TCTCAACAGAGAGATTCAGGCAATTAATCTTGTGCAATGCGGTGGAATGAGGCTTATCA 2172
Db 1929 TCTCAACAGAGAGATTCAGGCAATTAATCTTGTGCAATGCGGTGGAATGAGGCTTATCA 1988
QY 2173 AACTCTTCTTAAGGTAACCTGGAAGTCAATTTGACAGAGACTTTGGAAGAACTTTTCA 2232
Db 1989 AACTCTTCTTAAGGTAACCTGGAAGTCAATTTGACAGAGACTTTGGAAGAACTTTTCA 2048
QY 2233 TAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2292
Db 2049 TAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2108
QY 2293 CCAAGAGTGTGATGACAGAGCTTATGACAGCTCATCAACCAACCCATCTCAACAGAT 2352
Db 2109 CCAAGAGTGTGATGACAGAGCTTATGACAGCTCATCAACCAACCCATCTCAACAGAT 2168
QY 2353 GGATGATGTTCTGTGAACAGTTTGAAGAGGAGCCGAAACACAGTGTGAGAGGCAAG 2412

Db	2169	GGATGAGTTCTGTGAACAAGTTGGAAAAAGGACCGAAAAACA	CTGGCAAGGCCAGG	2228
Oy	2413	ACATACCCGAGGGAACAGTAAACAATGGAAGCACTTACAGA	AAATATGAAAAGSTAGAA	2472
Db	2229	ACATACCCGAGGGAACAGTAAACAATGGAAGCACTTACAGA	AAATATGAAAAGSTAGAA	2288
Oy	2473	CAGAGAGACCCACGAATTTGAGAGGGACCCAGAGATGTCTG	AGCTGCATTAACCTTGA	2532
Db	2289	CAGAGAGACCCACGAATTTGAGAGGGACCCAGAGATGTCTG	AGCTGCATTAACCTTGA	2348
Oy	2533	AACTCAAAACAAGTAGAAATCTTGCTAGCAATTA	CTGGAAAACAATGCAATATACAT	2592
Db	2349	AACTCAAAACAAGTAGAAATCTTGCTAGCAATTA	CTGGAAAACAATGCAATATACAT	2408
Oy	2593	GAACTTTTTTCAATGATATGATGTGATGTTTACAATG	TGTGGAAATTCAGCTGAGTTCCA	2652
Db	2409	GAACTTTTTTCAATGATATGATGTGATGTTTACAATG	TGTGGAAATTCAGCTGAGTTCCA	2468
Oy	2653	CCAAATTAATAATTAATCAATGAGTAACCTTCTTAAT	TAGGCTTTTTTCTTAATACC	2709
Db	2469	CCAAATTAATAATTAATCAATGAGTAACCTTCTTAAT	TAGGCTTTTTTCTTAATACC	2525

```

RESULT 5
PCT-US02-07826--283
; Sequence 283, Application PC/TU0207826
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer
; FILE REFERENCE: MRI-030PC
; CURRENT APPLICATION NUMBER: PCT/US02/07826
; CURRENT FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 283
; LENGTH: 2530
; TYPE: DNA
; ORGANISM: Homo sapiens
PCT-US02-07826--283

Query Match          92.6%; Score 2508; DB 1; Length 2530;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      193  CTGCAGATGGCGCTGTTACTAGAGATTGCTGTCCTTTCTGGGAGATTACTTAGAGC 252
DB      9    CTGCAGATGGCGCTGTTACTAGAGATTGCTGTCCTTTCTGGGAGATTACTTAGAGC 68
QY      253  AAGAGCAACATATCAGATGGGAAAGAACAAATGGCAAGGCTGAATAATTACTACAAGA 312
DB      69  AAGAGCAACATATCAGAAATGGGAAACAAATGGCAAGGCTGAATAATTACTACAAGA 128
QY      313  AATGTGGAATTCACAATATGATCACTTCAATGGCTTGGCCACAGCTCCAGTTATCA 372
DB      129  AATGTGGAATTCACAATATGATCACTTCAATGGCTTGGCCACAGCTCCAGTTATCA 188
QY      373  TACCTTCTTTTGGATGAGGAACGAGATGCGCTGATGTTGGAGCAAGGATCATAAT 432
DB      1    TACCTTCTTTTGGATGAGGAACGAGATGCGCTGATGTTGGAGCAAGGATCATAAT 188

```

Db	189	TACCTCTCTTTGGATGAGGAACGAGTAGGCTGATGTTGGAGCAAAAGATCAATAT	248
Qy	433	TTCAATGACCTGGTTAAATATCAAGGATTTTCAAAAGATGTGGCCAGTACTTAAAC	492
Db	249	TTCAATGACCTGGTTAAATATCAAGGATTTTCAAAAGATGTGGCCAGTACTTAAAC	308
Qy	493	CAGAAAGATGAAATGCAAGTGGGCTGGAAAAGACATCTTGAAAGAAATGTCTAATTTCA	552
Db	309	CAGAAAGATGAAATGCAAGTGGGCTGGAAAAGACATCTTGAAAGAAATGTCTAATTTCA	368
Qy	553	CAAGTACTTAAGGATATATACAGACTCATTTGACGCTGTGGAAAGGGGGCTTTTCA	612
Db	369	CAAGTACTTAAGGATATATACAGACTCATTTGACGCTGTGGAAAGGGGGCTTTTCA	428
Qy	613	TCCAATTTGACCTACATTTGAAATTTGACATCATCTTGAGACAAATATTTTAACTGGA	672
Db	429	TCCAATTTGACCTACATTTGAAATTTGACATCATCTTGAGACAAATATTTTAACTGGA	488
Qy	673	GAACTCACATTTTGAACGCGCCGTGGGAAGTCCATATGACCTTAAGCTGTCAAGC	732
Db	489	GAACTCACATTTTGAACGCGCCGTGGGAAGTCCATATGACCTTAAGCTGTCAAGC	548
Qy	733	ATCCCTTAAATAGATGAGAAATTAATCTGGAACCTGCACTGATTTTATGGGGCGAGA	792
Db	549	ATCCCTTAAATAGATGAGAAATTAATCTGGAACCTGCACTGATTTTATGGGGCGAGA	608
Qy	793	CTTGTCTATCTTCCGAACCTCTTGGGCAACAACCCATTCAGACAGACATGATTC	852
Db	609	CTTGTCTATCTTCCGAACCTCTTGGGCAACAACCCATTCAGACAGACATGATTC	668
Qy	853	CAGGTGCTCAATGATCCAAAGTTCTATTAGTCCCACTCATCTCAGAGATGACAAATCC	912
Db	669	CAGGTGCTCAATGATCCAAAGTTCTATTAGTCCCACTCATCTCAGAGATGACAAATCC	728
Qy	913	TGAAGTAGCAAAAGTATCTTTTCTCCCGGAAAATGCAATAGATGGAGAACTCTGG	972
Db	729	TGAAGTAGCAAAAGTATCTTTTCTCCCGGAAAATGCAATAGATGGAGAACTCTGG	788
Qy	973	AAAAGCTACTACGCTAGATAGGTCAGATATGCAAGATACCTTTGGAGGGCACAGAA	1032
Db	789	AAAAGCTACTACGCTAGATAGGTCAGATATGCAAGATACCTTTGGAGGGCACAGAA	848
Qy	1033	TCTGTGTAATAATGAGCAACATTCCTCAAAAGCTGCTGATTTGCTCAGTGCAGGTCC	1092
Db	849	TCTGTGTAATAATGAGCAACATTCCTCAAAAGCTGCTGATTTGCTCAGTGCAGGTCC	908
Qy	1093	AAATGGCATTAACACTATTTTGATGAATCTGAGATGTATTTCTTAATGAACCTTAAAG	1152
Db	909	AAATGGCATTAACACTATTTTGATGAATCTGAGATGTATTTCTTAATGAACCTTAAAG	968
Qy	1153	TCCTAAAAATCAGTTGATATGAGAGTATTACGACTCCAGATACATTTTCAAGAGATC	1212
Db	969	TCCTAAAAATCAGTTGATATGAGAGTATTACGACTCCAGATACATTTTCAAGAGATC	1022
Qy	1213	AGCCGTGTATATGATATGACATGATGTGAGAAAGGTGTTCTTGATCATATGCCA	1272
Db	1029	AGCCGTGTATATGATATGACATGATGTGAGAAAGGTGTTCTTGATCATATGCCA	1088
Qy	1273	CAGGATGGAACCACTATCAATGGGTGCTTATTCAGAAAGAGTCCCTTATCCAGGCTC	1332
Db	1089	CAGGATGGAACCACTATCAATGGGTGCTTATTCAGAAAGAGTCCCTTATCCAGGCTC	1148
Qy	1333	AGGAACCTGTGCCAGCAAAACATTTGGGTGTTTGAACCTTCAAAAGACCTTCTGATGA	1392
Db	1149	AGGAACCTGTGCCAGCAAAACATTTGGGTGTTTGAACCTTCAAAAGACCTTCTGATGA	1208
Qy	1393	TGTTATAACCTTTGCAAGAGTCAATCCAGCCATGACATCCAGTGTTTCTATGAACAA	1452
Db	1209	TGTTATAACCTTTGCAAGAGTCAATCCAGCCATGACATCCAGTGTTTCTATGAACAA	1268
Qy	1453	TCGCCCAATAGTATCAAAACGATGTAAATTAATTCATTAACAAATTTGTGATGACCG	1512
Db	1269	TCGCCCAATAGTATCAAAACGATGTAAATTAATTCATTAACAAATTTGTGATGACCG	1328

APPLICANT: Karen GIATT
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
FILE REFERENCE: MRI-030
CURRENT APPLICATION NUMBER: US/10/097,340
CURRENT FILING DATE: 2002-03-14
PRIOR APPLICATION NUMBER: 60/276,025
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/325,149
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/276,026
PRIOR FILING DATE: 2001-03-14
PRIOR APPLICATION NUMBER: 60/324,967
PRIOR FILING DATE: 2001/09/26
PRIOR APPLICATION NUMBER: 60/311,732
PRIOR FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: 60/325,102
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/323,580
NUMBER OF SEQ ID NOS: 363
SOFTWARE: FastSeq for Windows Version 4.0
SPO ID NO 283
LENGTH: 2530
TYPE: DNA
ORGANISM: Homo sapiens
US-10-097-340-283

Query Match 92.6%; Score 2508; DB 45; Length 2530;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 133 CTGACGATGGGCTGTTACTAGAGTTGTCTCTTTCTGGGAGTATCTTACAGC 252
DB 9 CTGACGATGGGCTGTTACTAGAGTTGTCTCTTTCTGGGAGTATCTTACAGC 68
QY 253 AAGACCAACTATCAGATGGAGAGACATGTGCCAAGCTGAATATCTTACAAAG 312
DB 69 AAGACCAACTATCAGATGGAGAGACATGTGCCAAGCTGAATATCTTACAAAG 128
QY 313 AATGTTGAATCCAAATGATGATCACTTCAATGGCTTGGCAACAGCTCCAGTTATCA 372
DB 129 AATGTTGAATCCAAATGATGATCACTTCAATGGCTTGGCAACAGCTCCAGTTATCA 188
QY 373 TACCTTCTTTGGAG 432
DB 189 TACCTTCTTTGGAG 248
QY 433 TTCAATTCGACCTGTTAATATCAAGAGATTTTCAAAAGATTGTGTGCCAGTATCTTAC 492
DB 249 TTCAATTCGACCTGTTAATATCAAGAGATTTTCAAAAGATTGTGTGCCAGTATCTTAC 308
QY 493 CAGAGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 552
DB 309 CAGAGAGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 368
QY 553 CAAAGTACTTAAGGATATATCAAGCTCACTTGTACGCTGTGAGAGAGAGAGAGAG 612
DB 369 CAAAGTACTTAAGGATATATCAAGCTCACTTGTACGCTGTGAGAGAGAGAGAGAG 428
QY 613 TCCAAATTCGACCTAATTAATGAGATCATCTGAGAGCAATATTTTAACTGGA 672
DB 429 TCCAAATTCGACCTAATTAATGAGATCATCTGAGAGCAATATTTTAACTGGA 488
QY 673 GAATCTCACTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 732
DB 489 GAATCTCACTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 548
QY 733 ATCCCTTTAATAG 792
DB 549 ATCCCTTTAATAG 608
QY 793 CTTGTCTATCTTCGAACTCTTGGGACCAACCAATCAGAGAGAGAGAGAGATTC 852

DB 609 CTTGTCTATCTTCGAACTCTTGGGACACCAACCAATCAGAGAGAGAGAGATTC 668
QY 853 CAGGTGGCTCAATGATCCAAAGTTCATTAGGCCACCTCATCTCAGAGATGACATCC 912
DB 669 CAGGTGGCTCAATGATCCAAAGTTCATTAGGCCACCTCATCTCAGAGATGACATCC 728
QY 913 TGAAGATGACAAAGATATCTTTTCTCCGTAAGAGAGAGAGAGAGAGAGAGAGAG 972
DB 729 TGAAGATGACAAAGATATCTTTTCTCCGTAAGAGAGAGAGAGAGAGAGAGAGAG 788
QY 973 AAAAGCTTCTCAAGCTGAGATAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 1032
DB 789 AAAAGCTTCTCAAGCTGAGATAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 848
QY 1033 TGTGTGATTAATAG 1092
DB 849 TGTGTGATTAATAG 908
QY 1093 AAATGGCAATGACATCTATTTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1152
DB 909 AAATGGCAATGACATCTATTTTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAG 968
QY 1153 TCTTAAATATCAAGTTGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1212
DB 969 TCTTAAATATCAAGTTGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 1028
QY 1213 AGCGGTGTATGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1272
DB 1029 AGCGGTGTATGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1088
QY 1273 CAGGATGAG 1332
DB 1089 CAGGATGAG 1148
QY 1333 AGGAACTTGTCCAG 1392
DB 1149 AGGAACTTGTCCAG 1208
QY 1393 TGTATTAATCTTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1452
DB 1209 TGTATTAATCTTTGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1268
QY 1453 TCGCCCAATGATGATCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1512
DB 1269 TCGCCCAATGATGATCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1328
QY 1513 AGTGAATGACAG 1572
DB 1329 AGTGAATGACAG 1388
QY 1573 TCTTAAAGTATGATCAATCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1632
DB 1389 TCTTAAAGTATGATCAATCTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1448
QY 1633 AGAATGACAGATTTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1692
DB 1449 AGAATGACAGATTTTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1508
QY 1693 GCAACCAATATATATGTTGTTCAACGGCTGGGAGTCCAGCTTCAACCGGTGGA 1752
DB 1509 GCAACCAATATATATGTTGTTCAACGGCTGGGAGTCCAGCTTCAACCGGTGGA 1568
QY 1753 TATTTAAGGAG 1812
DB 1569 TATTTAAGGAG 1628
QY 1813 TGGTTCTGATGATCTGCTATTTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1872
DB 1629 TGGTTCTGATGATCTGCTATTTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1688
QY 1873 AAGAAATGAG 1932

MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 16..2331
 US-08-488-935-1

Query Match 92.6%; Score 2508; DB 9; Length 2601;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	193	CTGCACATGGGCTGGTAACTAGATGCTGCTCTTTCTGGGAGTATTACTTACAGC	252
DB	9	CTGCACATGGGCTGGTAACTAGATGCTGCTCTTTCTGGGAGTATTACTTACAGC	68
QY	253	AAGAGCAACTATCAGAAATGGAAGAACATGTGCCAGGCTGAATTTATCTACAAAG	312
DB	69	AAGAGCAACTATCAGAAATGGAAGAACATGTGCCAGGCTGAATTTATCTACAAAG	128
QY	313	AATGTTGAATCCAAATGTGATCACTTTCAATGGCTGGCCAAAGCTCCAGTTATCA	372
DB	129	AATGTTGAATCCAAATGTGATCACTTTCAATGGCTGGCCAAAGCTCCAGTTATCA	188
QY	373	TACCTTCCTTTGGAGAGAGAGAGTGTGTATGTTGGAGCAAGATCACTATT	432
DB	189	TACCTTCCTTTGGAGAGAGAGTGTGTATGTTGGAGCAAGATCACTATT	248
QY	433	TTCAATCGACCTGTTAATATCAAGGATTTTCAAAAGATTGTGGCCAGTATCTTACAC	492
DB	249	TTCAATCGACCTGTTAATATCAAGGATTTTCAAAAGATTGTGGCCAGTATCTTACAC	308
QY	493	CAGAGAGATGAATGCAAGTGGGCTGGAAAAGACATCTGAAAGAAATGTCTAATTCAT	552
DB	309	CAGAGAGATGAATGCAAGTGGGCTGGAAAAGACATCTGAAAGAAATGTCTAATTCAT	368
QY	553	CAAGGACTTAAGGCATATATATCACTGATGCTGTGAGCAAGGAGGCTTTTCA	612
DB	369	CAAGGACTTAAGGCATATATATCACTGATGCTGTGAGCAAGGAGGCTTTTCA	428
QY	613	TCCAAATTTGACCTTACATTTGAATTTGACATCATCTGAGAGCAATATTTTAAAGCTGGA	672
DB	429	TCCAAATTTGACCTTACATTTGAATTTGACATCATCTGAGAGCAATATTTTAAAGCTGGA	488
QY	673	GAACTCACATTTTGAAGACGGCCGTGGAGAGAGTCCATATGACCTTAAGCTGTGACAC	732
DB	489	GAACTCACATTTTGAAGACGGCCGTGGAGAGAGTCCATATGACCTTAAGCTGTGACAC	548
QY	733	ATCCCTTTAATATGATGAGATTAATCTGTGAACTGACAGCTGATTTTATGGGCGAGA	792
DB	549	ATCCCTTTAATATGATGAGATTAATCTGTGAACTGACAGCTGATTTTATGGGCGAGA	608
QY	793	CTTTGCTATCTTCCGAACTCTTGGGACCAACCCCATCAGAGACAGAGATGATTC	852
DB	609	CTTTGCTATCTTCCGAACTCTTGGGACCAACCCCATCAGAGACAGAGATGATTC	668
QY	853	CAGGTGCTCAATGATCCAAAGTCTTAAGTCCCACTCATCTCAGAGATGACATTC	912
DB	669	CAGGTGCTCAATGATCCAAAGTCTTAAGTCCCACTCATCTCAGAGATGACATTC	728
QY	913	TGAAGATGACAAAGTATCTTTTCTTCCGTGAAATGCAATATGATGAGAACACTCTGG	972
DB	729	TGAAGATGACAAAGTATCTTTTCTTCCGTGAAATGCAATATGATGAGAACACTCTGG	788
QY	973	AAAAGCTACTCAGCTAGATAGTCAATATGCAAGATGACTTTGAGGGGACAGAG	1032
DB	789	AAAAGCTACTCAGCTAGATAGTCAATATGCAAGATGACTTTGAGGGGACAGAG	848
QY	1033	TCTGTGATTAATGACAACTTCTCAAGCTGTCTGATTTGTCTCAGTCCAGTTC	1092
DB	849	TCTGTGATTAATGACAACTTCTCAAGCTGTCTGATTTGTCTCAGTCCAGTTC	908
QY	1093	AAATGCAATGACATCTATTTTGAATGATGACAGATGATTTCTTAATGAATTTAAAGA	1152
DB	909	AAATGCAATGACATCTATTTTGAATGATGACAGATGATTTCTTAATGAATTTAAAGA	968

QY	1153	TCTTAAAAATCCAGTGTATATGAGTGTTTACGACTTCCAGTAACTTTCAAGGATC	1212
DB	969	TCTTAAAAATCCAGTGTATATGAGTGTTTACGACTTCCAGTAACTTTCAAGGATC	1028
QY	1213	ACCCTGTGTATATGATAGATGATGTGAGAAAGGTTCTTGGTCCATATGCCA	1272
DB	1029	ACCCTGTGTATATGATAGATGATGTGAGAAAGGTTCTTGGTCCATATGCCA	1088
QY	1273	CAGGATGAGACCCACTATCAATGGTGCCCTTATCAAGAAAGTCCCTATCCAGGCC	1332
DB	1089	CAGGATGAGACCCACTATCAATGGTGCCCTTATCAAGAAAGTCCCTATCCAGGCC	1148
QY	1333	AGGAATGTCCTCCACCAAAACATTTGGTGTGTTTACTCTACAAAGACCTTCGATGA	1392
DB	1149	AGGAATGTCCTCCACCAAAACATTTGGTGTGTTTACTCTACAAAGACCTTCGATGA	1208
QY	1393	TGTTATTAACCTTTGCAAGAGTCAATCCAGATGATCAATCCAGTTCCTATGAACA	1452
DB	1209	TGTTATTAACCTTTGCAAGAGTCAATCCAGATGATCAATCCAGTTCCTATGAACA	1268
QY	1453	TGCCCCATATGATCAAAAACGAGATGTAATTAATTAATTAACAAATTTGCTGACCG	1512
DB	1269	TGCCCCATATGATCAAAAACGAGATGTAATTAATTAATTAACAAATTTGCTGACCG	1328
QY	1513	AGTGAATGCAAGATGAGACAGTATGATTTATGTTATCGGAACAGATGTTGGAGCGT	1572
DB	1329	AGTGAATGCAAGATGAGACAGTATGATTTATGTTATCGGAACAGATGTTGGAGCGT	1388
QY	1573	TCTTAAAGTATGTTCAATTCCTAAGAGACTTGGTATGATTTAGAAAGTTCGTGCGA	1632
DB	1389	TCTTAAAGTATGTTCAATTCCTAAGAGACTTGGTATGATTTAGAAAGTTCGTGCGA	1448
QY	1633	AGAAATGACAGTTTTCGGGAAACCGACGTCATTTACCAATGAGCTTCCACTAAACA	1692
DB	1449	AGAAATGACAGTTTTCGGGAAACCGACGTCATTTACCAATGAGCTTCCACTAAACA	1508
QY	1693	GCAACACTATATATTTGTTCAACGAGCTGGGGTTGCCAGCTCCCTTACACCGGTGGA	1752
DB	1509	GCAACACTATATATTTGTTCAACGAGCTGGGGTTGCCAGCTCCCTTACACCGGTGGA	1568
QY	1753	TATTTACGGAAAGCGTGTGATGAGTGTGCTGCGCCGAGAACCTTACTGTGTTGGGA	1812
DB	1569	TATTTACGGAAAGCGTGTGATGAGTGTGCTGCGCCGAGAACCTTACTGTGTTGGGA	1628
QY	1813	TGTTCTGACATGTTCTGCTATTTTCCCACTGCAAGAGACGCAACAGACAAAGATAT	1872
DB	1629	TGTTCTGACATGTTCTGCTATTTTCCCACTGCAAGAGACGCAACAGACAAAGATAT	1688
QY	1873	AAGAAATGAGAACCACTGACTCATCTGTTACAGCTTACACCATATATACCAATGGCCA	1932
DB	1689	AAGAAATGAGAACCACTGACTCATCTGTTACAGCTTACACCATATATACCAATGGCCA	1748
QY	1933	CAGCCCTGAAGAGAAATCATATGATGTGAGAAATGATGATGATGATGATGATGATG	1992
DB	1749	CAGCCCTGAAGAGAAATCATATGATGTGAGAAATGATGATGATGATGATGATGATG	1808
QY	1993	TCCGAAGTCCGAGAGAGCGCTGTCTATTTGGCAATTCAGAGGGGAAATGAAAGCGGAA	2052
DB	1809	TCCGAAGTCCGAGAGAGCGCTGTCTATTTGGCAATTCAGAGGGGAAATGAAAGCGGAA	1868
QY	2053	AGAAAGATGACAGTGTATGATCATATCATCAGAGACAGATCAAGGCTTCTGCTACGTAG	2112
DB	1869	AGAAAGATGACAGTGTATGATCATATCATCAGAGACAGATCAAGGCTTCTGCTACGTAG	1928
QY	2113	TCTACACAGAGAGATTCAGGCAATTAACCTGTCATGCGGTGAGACATGGGTTCAATCA	2172
DB	1929	TCTACACAGAGAGATTCAGGCAATTAACCTGTCATGCGGTGAGACATGGGTTCAATCA	1988
QY	2173	AACTCTTTTAAAGTAAACCTTGAAGTCAATTTGAACAGAGATTTTGAAGAACTTTCTCA	2232
DB	1989	AACTCTTTTAAAGTAAACCTTGAAGTCAATTTGAACAGAGATTTTGAAGAACTTTCTCA	2048

QY 2233 TAAAGTGAATGATGAGATGCTCTAAGACCAAGAAATGTCCTAATAGCATGACCTAG 2292
DB 2049 TAAAGATGATGATGAGATGCTCTAAGACCAAGAAATGTCCTAATAGCATGACCTAG 2108
QY 2293 CCAGAGGCTGCTGATCAGAGACTTCAATGAGCTCAATCCACCCCAATCTCAACAGAT 2352
DB 2109 CCAGAGGCTGCTGATCAGAGACTTCAATGAGCTCAATCCACCCCAATCTCAACAGAT 2168
QY 2353 GGATGATGCTGCTGATCAGAGACTTCAATGAGCTCAATCCACCCCAATCTCAACAGAT 2412
DB 2169 GGATGATGCTGCTGATCAGAGACTTCAATGAGCTCAATCCACCCCAATCTCAACAGAT 2228
QY 2413 ACATACCCAGGGAACAGTAACTGAAGCACTTCAAGAAATTAAGAAAGGTAGAA 2472
DB 2229 ACATACCCAGGGAACAGTAACTGAAGCACTTCAAGAAATTAAGAAAGGTAGAA 2288
QY 2473 CAGGAGACCCAGGGAATTTGAGAGGCAACCCAGAGCTGCTGAGCTGATTAACCTTGA 2532
DB 2289 CAGGAGACCCAGGGAATTTGAGAGGCAACCCAGAGCTGCTGAGCTGATTAACCTTGA 2348
QY 2533 AACCTCAAAAGTGAAGAACTGCTAGACAAATACTGGAAGAAACAAATGCAATATACAT 2592
DB 2349 AACCTCAAAAGTGAAGAACTGCTAGACAAATACTGGAAGAAACAAATGCAATATACAT 2408
QY 2593 GAACCTTTTTCATGACATTAATGATGATGTTTCAATGCTGGAATTTGAGCTGATTCOA 2652
DB 2409 GAACCTTTTTCATGACATTAATGATGATGTTTCAATGCTGGAATTTGAGCTGATTCOA 2468
QY 2653 CCAATTATTAATTAATCATGATGATTAACCTTCTTAATAGGCTTTT 2700
DB 2469 CCAATTATTAATTAATCATGATGATTAACCTTCTTAATAGGCTTTT 2516

RESULT 10
US-08-489-057-1

Sequence 1, Application US/08489057
GENERAL INFORMATION:
APPLICANT: Goodman, Corey S.
APPLICANT: Kolodkin, Alex L.
APPLICANT: Matthes, David
APPLICANT: Bentley, David R.
APPLICANT: O'Connor, Timothy
TITLE OF INVENTION: Semaphorins
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: RICHARD ARON OSMAN, Ph.D.
STREET: 268 Bush Street, Suite 3200
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/489,057
FILING DATE: 09-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: B94-002-2/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2601 base pairs
TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 16..2331
US-08-489-057-1

Query Match 92.6%; Score 2508; DB 9; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 193 CTGACGATGGGCTGTTAACTAGATGCTGCTTTTCTGGGAGTATTACTTAACGC 252
DB 9 CTGACGATGGGCTGTTAACTAGATGCTGCTTTTCTGGGAGTATTACTTAACGC 68
QY 253 AAGAGCAAACTATGAGAAATGGGAGAAACATATGSCCAAGCTGAAATTAATCTCAAGA 312
DB 69 AAGAGCAAACTATGAGAAATGGGAGAAACATATGSCCAAGCTGAAATTAATCTCAAGA 128
QY 313 AATGTGGAATCAACAATGATGATCACTTCAATGGCTGGCCAAAGCTTCAGTTATCA 372
DB 129 AATGTGGAATCAACAATGATGATCACTTCAATGGCTGGCCAAAGCTTCAGTTATCA 188
QY 373 TACCTTCCTTTGGATGAGAAACGAGTAGGCTGTATGTTGACCAAGATCAATATT 432
DB 189 TACCTTCCTTTGGATGAGAAACGAGTAGGCTGTATGTTGACCAAGATCAATATT 248
QY 433 TTCAATGACCTGTTAATATCAAGATTTTCAAAAGATTGTGTGCGCAATTTTACAC 492
DB 249 TTCAATGACCTGTTAATATCAAGATTTTCAAAAGATTGTGTGCGCAATTTTACAC 308
QY 493 CAGAGAGATGAATGCAAGTGGGCTGGGAAAGACATGCTGAAAGATGTGTAATTTTCA 552
DB 309 CAGAGAGATGAATGCAAGTGGGCTGGGAAAGACATGCTGAAAGATGTGTAATTTTCA 368
QY 553 CAAGTAATTAAAGCATATATCAACTCACTTGTACGCTGTGAAACGGGGCTTTTCA 612
DB 369 CAAGTAATTAAAGCATATATCAACTCACTTGTACGCTGTGAAACGGGGCTTTTCA 428
QY 613 TCCAATTTGCACTTCAATTTGAATTTGACATCACTCTGAGGACAAATTTTAAAGCTGA 672
DB 429 TCCAATTTGCACTTCAATTTGAATTTGACATCACTCTGAGGACAAATTTTAAAGCTGA 488
QY 673 GAACCTCAATTTTGAAGACGGGCTGGGAAAGATGCAATGACCTTAAGCTGACAGC 732
DB 489 GAACCTCAATTTTGAAGACGGGCTGGGAAAGATGCAATGACCTTAAGCTGACAGC 548
QY 733 ATCCCTTTTAATAGATGAGAAATTAATCTGGAATGCACTGATTTTATGGGCGAGA 792
DB 549 ATCCCTTTTAATAGATGAGAAATTAATCTGGAATGCACTGATTTTATGGGCGAGA 608
QY 793 CTTTGTATCTTCCGAACCTTGGGCAACACCCCAATCAAGAGAGATGATTC 852
DB 609 CTTTGTATCTTCCGAACCTTGGGCAACACCCCAATCAAGAGAGATGATTC 668
QY 853 CAGGTGCTCAATATCAAGTTCATTAAGGCCACCTCATCTCAAGAGTGAATTC 912
DB 669 CAGGTGCTCAATATCAAGTTCATTAAGGCCACCTCATCTCAAGAGTGAATTC 728
QY 913 TGAAGATGCAAAAGTATCTTTTCTCGTGAAGATGCAATAGATGAGAAACATCTGG 972
DB 729 TGAAGATGCAAAAGTATCTTTTCTCGTGAAGATGCAATAGATGAGAAACATCTGG 788
QY 973 AAAAGCTACTGACCTTGAATAGTCAATGATGCAAGATGACTTTGAGAGGCAAGAG 1032
DB 789 AAAAGCTACTGACCTTGAATAGTCAATGATGCAAGATGACTTTGAGAGGCAAGAG 848
QY 1033 TCTGTGATTAATGAACAATCTTCAAGCTGCTGATTTGCTCAGTGCAGGCTCC 1092
DB 849 TCTGTGATTAATGAACAATCTTCAAGCTGCTGATTTGCTCAGTGCAGGCTCC 908
QY 1093 AAATGCAATTGACATCATTTTGAATGAACTGACAGATGATTTCTTAATGAATTTAAGA 1152

Db 909 AATGACATGACATCTATTTGATGAGACAGAGATGATCTTAAAGAACTTTAAAGA 968
 Qy 1153 TCTTAAAAATCCAGTTGATATGAGAGTTTACAGCTCCAGTAACTTTTCAAGGATC 1212
 Db 969 TCTTAAAAATCCAGTTGATATGAGAGTTTACAGCTCCAGTAACTTTTCAAGGATC 1028
 Qy 1213 AGCCGTGTATATGATAGATGAGATGATGAGAGAGGTTTCTTGGTCCATATGCCCA 1272
 Db 1029 AGCCGTGTATATGATAGATGAGATGATGAGAGAGGTTTCTTGGTCCATATGCCCA 1088
 Qy 1273 CAGGATGAGACCAACTATCAATGAGTCTTATCAAGAGAGTCCCTATCCAGGAC 1332
 Db 1089 CAGGATGAGACCAACTATCAATGAGTCTTATCAAGAGAGTCCCTATCCAGGAC 1148
 Qy 1333 AGGAACTTGTCCAGCAAAACATTTGGTGTGATCTCAAAAGAGACCTTCTGATGA 1392
 Db 1149 AGGAACTTGTCCAGCAAAACATTTGGTGTGATCTCAAAAGAGACCTTCTGATGA 1208
 Qy 1393 TGTATACCTTTGCAAGAGATCATCCAGCATGTAACATCCAGTGTCTTATGAACA 1452
 Db 1209 TGTATACCTTTGCAAGAGATCATCCAGCATGTAACATCCAGTGTCTTATGAACA 1268
 Qy 1453 TGGCCCAATATGATATCAAAAGATGATTAATTAATTAATTAATTAATTAATTAATTA 1512
 Db 1269 TGGCCCAATATGATATCAAAAGATGATTAATTAATTAATTAATTAATTAATTAATTA 1328
 Qy 1513 AGTGAATGAGAGATGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1572
 Db 1329 AGTGAATGAGAGATGAGACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1388
 Qy 1573 TCTTAAAGATGATTCATATCTTAAGAGACCTTGTATGATTAAGAGAGGTTCTGCTGA 1632
 Db 1389 TCTTAAAGATGATTCATATCTTAAGAGACCTTGTATGATTAAGAGAGGTTCTGCTGA 1448
 Qy 1633 AGAATATGACAGTTTTCGAGAACCCGACTGATATTAAGAGAGAGAGAGAGAGAGAGAGAG 1692
 Db 1449 AGAATATGACAGTTTTCGAGAACCCGACTGATATTAAGAGAGAGAGAGAGAGAGAGAGAG 1508
 Qy 1693 GCAACAATATATATGATGATCAAGGCTGGGTTGCGCAGCTCCCTTACACCGGATGA 1752
 Db 1509 GCAACAATATATATGATGATCAAGGCTGGGTTGCGCAGCTCCCTTACACCGGATGA 1568
 Qy 1753 TATTTACGGGAAAGCGTGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1812
 Db 1569 TATTTACGGGAAAGCGTGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1628
 Qy 1813 TGGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1872
 Db 1629 TGGTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1688
 Qy 1873 AAGAAATGAGAGACCACTGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1932
 Db 1689 AAGAAATGAGAGACCACTGATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1748
 Qy 1933 CAGCCCTGAGAGAGATCATCTATGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAG 1992
 Db 1749 CAGCCCTGAGAGAGATCATCTATGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAG 1808
 Qy 1993 TCCGAGTGCAGAGAGCGCTGCTGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAG 2052
 Db 1809 TCCGAGTGCAGAGAGCGCTGCTGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAG 1868
 Qy 2053 AGAAGATGAGAGATGATCATATCATCATCATCATCATCATCATCATCATCATCATCATCAT 2112
 Db 1869 AGAAGATGAGAGATGATCATATCATCATCATCATCATCATCATCATCATCATCATCATCAT 1928
 Qy 2113 TCTTAAACAGAGATGATGAGAGATGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAG 2172
 Db 1929 TCTTAAACAGAGATGATGAGAGATGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAG 1988
 Qy 2173 AACTCTTTTAAAGTAACTTGAAGTCAATGACAGAGAGATGAGAGAGATGAGAGAGATGAGAG 2232
 Db 1989 AACTCTTTTAAAGTAACTTGAAGTCAATGACAGAGAGATGAGAGAGATGAGAGAGATGAGAG 2048

Qy 2233 TAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2292
 Db 2049 TAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2108
 Qy 2293 CCAAGAGTCTGCTGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2352
 Db 2109 CCAAGAGTCTGCTGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 2168
 Qy 2353 GAT 2412
 Db 2169 GAT 2228
 Qy 2413 ACATACCCAGGAGACAGTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 2472
 Db 2229 ACATACCCAGGAGACAGTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 2288
 Qy 2473 CAGAGAGACCAAGATTTGAG 2532
 Db 2289 CAGAGAGACCAAGATTTGAG 2348
 Qy 2533 AACCTCAACAGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2592
 Db 2349 AACCTCAACAGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2408
 Qy 2593 GAACTTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2652
 Db 2409 GAACTTTTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2468
 Qy 2653 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2700
 Db 2469 CCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2516

RESULT 11

US-10-067-632-53

; Sequence 53, Application US/10067632

GENERAL INFORMATION:

APPLICANT: Goodman, Corey S.

Kolodkin, Alex L.

Methes, David R.

Bentley, David R.

O'Connor, Timothy

TITLE OF INVENTION: The Semaphorin Gene Family

NUMBER OF SEQUENCES: 100

CORRESPONDENCE ADDRESSES:

ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

STREET: 268 Bush Street, Suite 3200

CITY: San Francisco

STATE: CA

COUNTRY: USA

ZIP: 94104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/067,632

FILING DATE: 04-Feb-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/09/060,610

FILING DATE: <Unknown>

APPLICATION NUMBER: 08/835,268

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Oseman, Richard A.

REGISTRATION NUMBER: 36,627

REFERENCE/DOCKET NUMBER: B94-002-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 343-4341

TELEFAX: (415) 343-4342

TELEX: <Unknown>
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 2601 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 16..2331
SEQUENCE DESCRIPTION: SEQ ID NO: 53:
US-10-067-632-53
Query Match 92.6%; Score 2508; DB 45; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 193 CTGAGCATGGGCTGGTAACTAGAGATTGCTCTTTCTTGGGAGATTAATTACAGC 252
DB 9 CTGAGCATGGGCTGGTAACTAGAGATTGCTCTTTCTTGGGAGATTAATTACAGC 68
QY 253 AAGACCAACTATCGAATGGGAGAGAACATGTCAGGCTGAATTAATCTTACAAAGA 312
DB 69 AAGACCAACTATCGAATGGGAGAGAACATGTCAGGCTGAATTAATCTTACAAAGA 128
QY 313 AATGTGGAACTCAACATGTGATCACTTCAATGGCTGGCCAAAGCTCCAGTTATCA 372
DB 129 AATGTGGAACTCAACATGTGATCACTTCAATGGCTGGCCAAAGCTCCAGTTATCA 188
QY 373 TACCTTCTTTGGATGAGGAAAGAGAGAGTGTATGTTGGAGCAAGATCAGATATT 432
DB 189 TACCTTCTTTGGATGAGGAAAGAGAGAGTGTATGTTGGAGCAAGATCAGATATT 248
QY 433 TTCAATTCAGCCGTGTTAATATCAAGATTTTCAAAAGATTGTGTGCCAGTATCTTAC 492
DB 249 TTCAATTCAGCCGTGTTAATATCAAGATTTTCAAAAGATTGTGTGCCAGTATCTTAC 308
QY 493 CAGAAGATGAAATCAAGTGGCTGGAAGAAAGACATCTGAAAGATGCTAATTTTCA 552
DB 309 CAGAAGATGAAATCAAGTGGCTGGAAGAAAGACATCTGAAAGATGCTAATTTTCA 368
QY 553 CAAGGTACTTAAAGCATTAATCAGACTCTTGAAGCTGTGAGCGGGGCTTTTCA 612
DB 369 CAAGGTACTTAAAGCATTAATCAGACTCTTGAAGCTGTGAGCGGGGCTTTTCA 428
QY 613 TTCAATTTGCACTTATGAAATTGAGCATATCTGAGAGCAATATTTTAAGCTGA 672
DB 429 TTCAATTTGCACTTATGAAATTGAGCATATCTGAGAGCAATATTTTAAGCTGA 488
QY 673 GAACCTCACTTTTGAAGCGGCGTGGGAGAGTGCATATGACCTTAAGCTGCTGACG 732
DB 489 GAACCTCACTTTTGAAGCGGCGTGGGAGAGTGCATATGACCTTAAGCTGCTGACG 548
QY 733 ATCCCTTTTAATAGATGAGATTAATTAATCTGGAAGCTGAGCTGATTTATGAGGCGAGA 792
DB 549 ATCCCTTTTAATAGATGAGATTAATTAATCTGGAAGCTGAGCTGATTTATGAGGCGAGA 608
QY 793 CTTCCTATCTTGGGAATCTTGGGACCAACCAATCAGACAGAGCATGATTC 852
DB 609 CTTCCTATCTTGGGAATCTTGGGACCAACCAATCAGACAGAGCATGATTC 668
QY 853 CAGGTGGCTCAATGATCCAAAGTCAATTAAGTGCACCTCATCTGAGAGTGAACAATCC 912
DB 669 CAGGTGGCTCAATGATCCAAAGTCAATTAAGTGCACCTCATCTGAGAGTGAACAATCC 728
QY 913 TGAAGATGCAAAAGTATATCTTTTCTTCGTGAAATGCAATAGATGAGACACTCTGG 972
DB 729 TGAAGATGCAAAAGTATATCTTTTCTTCGTGAAATGCAATAGATGAGACACTCTGG 788
QY 973 AAAAGCTACTCAGCTAGATAGGTGATATGCAAGATGACTTTGAGGGGCAAGAG 1032
DB 789 AAAAGCTACTCAGCTAGATAGGTGATATGCAAGATGACTTTGAGGGGCAAGAG 848

QY 1033 TCTGTGATTAATAGAGCAACATTCCTCAAGCTGCTGATTTGCTGAGTCCAGGCTCC 1092
DB 849 TCTGTGATTAATAGAGCAACATTCCTCAAGCTGCTGATTTGCTGAGTCCAGGCTCC 908
QY 1093 AATGCACTGACACTCTATTTGATGAATGCAAGATGATTTCTTAATGAATTTAAGA 1152
DB 909 AATGCACTGACACTCTATTTGATGAATGCAAGATGATTTCTTAATGAATTTAAGA 968
QY 1153 TCTTAATAATCAGTGTATATGAGTGTATAGACTTCCAGTAACTTTCAAGGGATC 1212
DB 969 TCTTAATAATCAGTGTATATGAGTGTATAGACTTCCAGTAACTTTCAAGGGATC 1028
QY 1213 AGCGGTGTATATGATAGCATGATGATGATGAGAGGCTGTTCTTGGTCCATATGCCA 1272
DB 1029 AGCGGTGTATATGATAGCATGATGATGAGAGGCTGTTCTTGGTCCATATGCCA 1088
QY 1273 CAGGATGAGCCCACTATCAATGGGTGCTTATCAAGAAAGTCCCTATCAAGGCT 1332
DB 1089 CAGGATGAGCCCACTATCAATGGGTGCTTATCAAGAAAGTCCCTATCAAGGCT 1148
QY 1333 AGGAACCTGTCCAGCAAAACATTTGGTGTGACTCTACAAAGACCTTCTGATGA 1392
DB 1149 AGGAACCTGTCCAGCAAAACATTTGGTGTGACTCTACAAAGACCTTCTGATGA 1208
QY 1393 TGTATAACCTTTGCAAGAAATGCTCCAGCATGTAACATCCAGTGTTCCTATGAACA 1452
DB 1209 TGTATAACCTTTGCAAGAAATGCTCCAGCATGTAACATCCAGTGTTCCTATGAACA 1268
QY 1453 TCGCCCAATAGTATCAAAACGAGTGAATTAATTAATCAAAATGTCGTAGACCG 1512
DB 1269 TCGCCCAATAGTATCAAAACGAGTGAATTAATTAATCAAAATGTCGTAGACCG 1328
QY 1513 AGTGAATGCAAAATGAGACGATAGATGATGATGATGATGATGATGATGATGATG 1572
DB 1329 AGTGAATGCAAAATGAGACGATAGATGATGATGATGATGATGATGATGATGATG 1388
QY 1573 TCTTAAGTATGTTCAATTCCTAAGAGACTTGGTATGATTTAAGAGTGTCTGCTGA 1632
DB 1389 TCTTAAGTATGTTCAATTCCTAAGAGACTTGGTATGATTTAAGAGTGTCTGCTGA 1448
QY 1633 AGAAATGACAGTTTTTGGGAAACGAGTGTATTTAGCAATGAGCTTCCACTAAGCA 1692
DB 1449 AGAAATGACAGTTTTTGGGAAACGAGTGTATTTAGCAATGAGCTTCCACTAAGCA 1508
QY 1693 GCAACAACTATATATGTTGTTCAACGCTGGGCTGCCAGCTCCCTTTACACGGGTGA 1752
DB 1509 GCAACAACTATATATGTTGTTCAACGCTGGGCTGCCAGCTCCCTTTACACGGGTGA 1568
QY 1753 TATTTAAGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGACCTTAAGCTGCTGGA 1812
DB 1569 TATTTAAGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGACCTTAAGCTGCTGGA 1628
QY 1813 TGGTTCGATGTTCTGCTATTTTCCCATGCAAGAGAGCAAGACGCAAGATAT 1872
DB 1629 TGGTTCGATGTTCTGCTATTTTCCCATGCAAGAGAGCAAGACGCAAGATAT 1688
QY 1873 AAGAAATGAGAACCCATGATCTGACTGTTGAGACTTACCATGATATATCACTGCGCA 1932
DB 1689 AAGAAATGAGAACCCATGATCTGACTGTTGAGACTTACCATGATATATCACTGCGCA 1748
QY 1933 CAGGCTGGAAGAGAAATCATCTATGAGTGTGAGAAATAGTGAACATTTTGAATGAG 1992
DB 1749 CAGGCTGGAAGAGAAATCATCTATGAGTGTGAGAAATAGTGAACATTTTGAATGAG 1808
QY 1993 TCCGAAGTCGAGAGAGGCTGTCTATTTGGCAATTCAGAGGGCAATGAAGCGGAA 2052
DB 1809 TCCGAAGTCGAGAGAGGCTGTCTATTTGGCAATTCAGAGGGCAATGAAGCGGAA 1868
QY 2053 AGAAGATGCAAGTGTATGATATATATATGAGAGAGATTAAGGCTTCTGCTACGTAG 2112
DB 1869 AGAAGATGCAAGTGTATGATATATATATGAGAGAGATTAAGGCTTCTGCTACGTAG 1928

Dd	2341	AGAAACTCAACAGTAGAATCTTGCTGACATATACTGGAAAAACAATTGCATATTA	2400
Qy	2590	CATGAACTTTTTCATGGCATTATGTGAGTGTTTACAAATGGTGGAATTCAGCTGAGTT	2649
Dd	2401	CATGAACTTTTTCATGGCATTTATGTGGATGTTTACAAATGGTGGAATTCAGCTGAGATT	2460
Qy	2650	CCACCAATTTAAATTTAAATCCATGAGAACTTTCCTAATATGGCTTTTTCCTAATACC	2709
Dd	2461	CCACCAATTTAAATTTAAATCAATGAGAACCTTTCCTAATATGGCTTTTTCCTAATACC	2520

```

1  GENERATED INFORMATION:
2  APPLICANT: Holtzman, Douglas A.
3  TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
4  TITLE OF INVENTION: THEREFOR
5  FILE REFERENCE: 1600, 1165, 001
6  CURRENT APPLICATION NUMBER: US/09/649,164
7  CURRENT FILING DATE: 2000-08-25
8  PRIOR APPLICATION NUMBER: 60/1150,606
9  PRIOR FILING DATE: 1999-08-25
10 NUMBER OF SEQ ID NOS: 9995
11 SOFTWARE: FastSeq for Windows Version 4.0
12 SEQ ID NO 8352
13 LENGTH: 2256
14 TYPE: DNA
15 ORGANISM: Homo sapiens
16 FEATURE:
17 NAME/KEY: misc feature
18 LOCATION: (1)..(2256)
19 OTHER INFORMATION: n = A,T,C or G
20 US-09-649-164-8352

```

QY	22	GTCTGCTTTTCGGGAGTATTACTTTCACGGAAGACAACTTCAGATTGGGAGAAC	28
Db	163	GTCTGCTTTTCGGGAGTATTACTTTCACGGAAGACAACTTCAGATTGGGAGAAC	222
QY	281	AATGTSCCAAGGCTGAATTATCTCAAGAAGATGTTGGAATCCAACTATGTATCACT	340
Db	223	AATGTSCCAAGGCTGAATTATCTCAAGAAGATGTTGGAATCCAACTATGTATCACT	282
QY	341	TTTCAATGGCTTGGCCCAACGCTCCAGTTATCATACCTCTTTTGGATGAGGAACGGAGT	400
Db	283	TTTCAATGGCTTGGCCCAACGCTCCAGTTATCATACCTCTTTTGGATGAGGAACGGAGT	342
QY	401	AGGCTTATGTTGGAGCAAGAGATACATATTTTCATTCGACCTGGTTAATATCAAGAT	460
Db	343	AGGCTTATGTTGGAGCAAGAGATACATATTTTCATTCGACCTGGTTAATATCAAGAT	402
QY	461	TTTCAAAAGATTGTGTGGCCAGTATCTTACACCGAAGAGATGATGCAAGTGGCTTGA	520
Db	403	TTTCAAAAGATTGTGTGGCCAGTATCTTACACCGAAGAGATGATGCAAGTGGCTTGA	462
QY	521	AAAGACATCTGTAAGAAGATGTGCTAATTTCAATCAAGATCTTAAGGCATATATCAACT	580
Db	463	AAAGACATCTGTAAGAAGATGTGCTAATTTCAATCAAGATCTTAAGGCATATATCAACT	522
QY	581	CACCTGTACGCTGTGGAACGGGGGCTTTTCATCCAAATTTGCACCTACATTTGAATTGGA	640
Db	523	CACCTGTACGCTGTGGAACGGGGGCTTTTCATCCAAATTTGCACCTACATTTGAATTGGA	582
QY	641	CATCATCTCGAGAGACATATTTTTTAAAGCTGGAGAACTCAATTTTGAAGAAACGGCCGTGGG	700
Db	583	CATCATCTCGAGAGACATATTTTTTAAAGCTGGAGAACTCAATTTTGAAGAAACGGCCGTGGG	642

QY 701 AAGAGTCATATGACCCCTAAGCTGCTGACAGACATCCCTTTTAAATAGATGAGATTAATAC 760
 Db 643 AAGAGTCATATGACCCCTAAGCTGCTGACAGACATCCCTTTTAAATAGATGAGATTAATAC 702
 QY 761 TCTGGAACCTGACGCTGATTTTATGAGGCGAGACCTTTTGTCTATCTTCCGAACCTTTGGGAC 820
 Db 703 TCTGGAACCTGACGCTGATTTTATGAGGCGAGACCTTTTGTCTATCTTCCGAACCTTTGGGAC 762
 QY 821 CACCAACCAATGACGACAGACAGATGATTCAGAGTGGCTCAATGATCCAAAGTTCATT 880
 Db 763 CACCAACCAATGACGACAGACAGATGATTCAGAGTGGCTCAATGATCCAAAGTTCATT 822
 QY 881 AGTGGCCACCTCATCTCAAGAGATGACAACTCTGAAGATGACAAAGTATATCTTTTCTTC 940
 Db 823 AGTGGCCACCTCATCTCAAGAGATGACAACTCTGAAGATGACAAAGTATATCTTTTCTTC 882
 QY 941 CGTGAATAATGCAATATGAGAGACCTCTGAAAAAGTATCTACGCTAGATAGGTGAG 1000
 Db 883 CGTGAATAATGCAATATGAGAGACCTCTGAAAAAGTATCTACGCTAGATAGGTGAG 942
 QY 1001 ATATGCAAGATGACCTTTGAGGGGACAGAGATCTGTGTAATTAATGACAAATCTCTC 1060
 Db 943 ATATGCAAGATGACCTTTGAGGGGACAGAGATCTGTGTAATTAATGACAAATCTCTC 1002
 QY 1061 AAAGCTGCTGATTTGCTCAAGTGCAGAGTCCAAATGCAATGACATCAATTTTGAAGA 1120
 Db 1003 AAAGCTGCTGATTTGCTCAAGTGCAGAGTCCAAATGCAATGACATCAATTTTGAAGA 1062
 QY 1121 CTGCAAGATGATTTCTTAATGACCTTTAAAGATCTTAAATCCCAATGATTAATGAGAG 1180
 Db 1063 CTGCAAGATGATTTCTTAATGACCTTTAAAGATCTTAAATCCCAATGATTAATGAGAG 1122
 QY 1181 TTACGACTTCAGATTAACATTTTCAAGGATCAGCCGTGTATATGATAGATGATGAT 1240
 Db 1123 TTACGACTTCAGATTAACATTTTCAAGGATCAGCCGTGTATATGATAGATGATGAT 1182
 QY 1241 GTGAGAGGGTGTCTTGTGTCATATATGCCCACAGGGATGACCCCAATATCAATGGAGT 1300
 Db 1183 GTGAGAGGGTGTCTTGTGTCATATATGCCCACAGGGATGACCCCAATATCAATGGAGT 1242
 QY 1301 CCTTATCAAGAGAGATCCCCCTATCCACGGCCAGGAACTTTGCCGCAAAACATTTGGT 1360
 Db 1243 CCTTATCAAGAGAGATCCCCCTATCCACGGCCAGGAACTTTGCCGCAAAACATTTGGT 1302
 QY 1361 GGTTTGACTCTACAAAGAGACCTTCTGATGATGATTAATACCTTTGCAAGAGTCAATCA 1420
 Db 1303 GGTTTGACTCTACAAAGAGACCTTCTGATGATGATTAATACCTTTGCAAGAGTCAATCA 1362
 QY 1421 GCCATGATCAATCCAGTGTCTTCCATGAACAAATCCGCCCAATAGTATCAAAACGAGATGA 1480
 Db 1363 GCCATGATCAATCCAGTGTCTTCCATGAACAAATCCGCCCAATAGTATCAAAACGAGATGA 1422
 QY 1481 AATTATCAATTTTACCAAAATGTCGAGACCGAGTGGATGACAGAGATGAGATGATGAT 1540
 Db 1423 AATTATCAATTTTACCAAAATGTCGAGACCGAGTGGATGACAGAGATGAGATGATGAT 1482
 QY 1541 GTTATGTTATGCAAGACAGATGTGGAGCCGTTCTTAAAGTATGATTTCAATCTTAAGAG 1600
 Db 1483 GTTATGTTATGCAAGACAGATGTGGAGCCGTTCTTAAAGTATGATTTCAATCTTAAGAG 1542
 QY 1601 ACTTGATA 1608
 Db 1543 ACTTGATA 1550

RESULT 14

US-09-652-816-7308
 ; Sequence 7308, Application US/09652816
 ; GENERAL INFORMATION:
 ; APPLICANT: Gutierrez-Ramos, Jose-Carlos
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
 ; FILE REFERENCE: 1600.1177-001

; CURRENT APPLICATION NUMBER: US/09/652,816
 ; CURRENT FILING DATE: 2000-08-31
 ; PRIOR APPLICATION NUMBER: 60/152,111
 ; PRIORITY FILING DATE: 1999-08-31
 ; NUMBER OF SEQ ID NOS: 9647
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 7308
 ; LENGTH: 2256
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(2256)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-652-816-7308

Query Match 51.2%; Score 1388; DB 28; Length 2256;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 221 GTCGTCTTTTCTGAGGAGTATTAATCTTACAGCAAGCAAACTATCAGAAATGGAAGAAC 280
 Db 163 GTCGTCTTTTCTGAGGAGTATTAATCTTACAGCAAGCAAACTATCAGAAATGGAAGAAC 222
 QY 281 AATGTCAGAGCTGAATTAATCTTACAAAGAAATGTTGAATCCAAATGTGATCACT 340
 Db 223 AATGTCAGAGCTGAATTAATCTTACAAAGAAATGTTGAATCCAAATGTGATCACT 282
 QY 341 TTCAATGCTTGGCCAAACAGCTCCAGTATATCAATCTTCTTTGAGAGAGAACGGAGT 400
 Db 283 TTCAATGCTTGGCCAAACAGCTCCAGTATATCAATCTTCTTTGAGAGAGAACGGAGT 342
 QY 401 AGCGTGTATGTTGAGCAAGAGATCAATATTTTCAATGCACTGCTGTTAATATCAAGAT 460
 Db 343 AGCGTGTATGTTGAGCAAGAGATCAATATTTTCAATGCACTGCTGTTAATATCAAGAT 402
 QY 461 TTTCAAAAGATTTGTGTGCGCAGATATCTTACACAGAGAGATGAATGCAAGTGGCTGGA 520
 Db 403 TTTCAAAAGATTTGTGTGCGCAGATATCTTACACAGAGAGATGAATGCAAGTGGCTGGA 462
 QY 521 AAAGCATCTGAAAGATATGCTTAATTTTCAATCAAGATCTTAAAGCATTAATCAAGT 580
 Db 463 AAAGCATCTGAAAGATATGCTTAATTTTCAATCAAGATCTTAAAGCATTAATCAAGT 522
 QY 581 CACTGTAGCGCTGTGGAAGCGGGGCTTTTATCAATTTTCAATCAATCAATCAATTTGGA 640
 Db 523 CACTGTAGCGCTGTGGAAGCGGGGCTTTTATCAATTTTCAATCAATCAATCAATTTGGA 582
 QY 641 CATCATCTGAGAGCAATATTTTAAAGCTGAGAACTCAATTTTGAAGACGGCCGTGGG 700
 Db 583 CATCATCTGAGAGCAATATTTTAAAGCTGAGAACTCAATTTTGAAGACGGCCGTGGG 642
 QY 701 AAGATCCATATGACCCCTAAGCTGTGACACATCCCTTAAATGAGAGAAATTAATAC 760
 Db 643 AAGATCCATATGACCCCTAAGCTGTGACACATCCCTTAAATGAGAGAAATTAATAC 702
 QY 761 TCTGGAACCTGACGCTGATTTTATGAGGCGAGACCTTTGCTATCTTCCGAACCTTTGGGAC 820
 Db 703 TCTGGAACCTGACGCTGATTTTATGAGGCGAGACCTTTGCTATCTTCCGAACCTTTGGGAC 762
 QY 821 CACCAACCAATGACGACAGACAGATGATTCAGAGTGGCTCAATGATCCAAAGTTCATT 880
 Db 763 CACCAACCAATGACGACAGACAGATGATTCAGAGTGGCTCAATGATCCAAAGTTCATT 822
 QY 881 AGTGGCCACCTCATCTCAAGAGATGACAACTCTGAAGATGACAAAGTATATCTTTTCTTC 940
 Db 823 AGTGGCCACCTCATCTCAAGAGATGACAACTCTGAAGATGACAAAGTATATCTTTTCTTC 882
 QY 941 CGTGAATAATGCAATATGAGAGACCTCTGAAAAAGTATCTACGCTAGATAGGTGAG 1000
 Db 883 CGTGAATAATGCAATATGAGAGACCTCTGAAAAAGTATCTACGCTAGATAGGTGAG 942
 QY 1001 ATATGCAAGATGACCTTTGAGGGGACAGAGATCTGTGTAATTAATGACAAATCTCTC 1060

Db 943 AATATGCAAGAAATGACTTTGGAGGACACAAAGCTGTGTAATAATGACCAATTCCTC 1002
Qy 1061 AAAGCTCGCTGATTTGCTCAGTGCAGAGTCCAAATGGCATGACACTATTTGATGA 1120
Db 1003 AAAGCTCGCTGATTTGCTCAGTGCAGAGTCCAAATGGCATGACACTATTTGATGA 1062
Qy 1121 CTGCAAGATGATATTCCTAATGAACTTTAAAGATCTTAAATTCAGTTGATATGAGAGT 1180
Db 1063 CTGCAAGATGATATTCCTAATGAACTTTAAAGATCTTAAATTCAGTTGATATGAGAGT 1122
Qy 1181 TTTAGACTTCCAGTAACATTTTCAAGGATCAGCCGTGTATGATATGATGATGAT 1240
Db 1123 TTTAGACTTCCAGTAACATTTTCAAGGATCAGCCGTGTATGATATGATGATGAT 1182
Qy 1241 GTGAGAGGGGTCTCTGTGTCATATGCCCCAGGATGACCCAACTATCAATGGGGT 1300
Db 1183 GTGAGAGGGGTCTCTGTGTCATATGCCCCAGGATGACCCAACTATCAATGGGGT 1242
Qy 1301 CCTATCAAGAAAGTCCCTATCCAGCCAGGAACTTGTCCAGCAAAACATTTGGT 1360
Db 1243 CCTATCAAGAAAGTCCCTATCCAGCCAGGAACTTGTCCAGCAAAACATTTGGT 1302
Qy 1361 GGTTTGACTCTACAAAGACCTTCTGATGATGTTATTAACCTTTGCAAGAACTATCA 1420
Db 1303 GGTTTGACTCTACAAAGACCTTCTGATGATGTTATTAACCTTTGCAAGAACTATCA 1362
Qy 1421 GCCATGTCATCCAGTGTCTTCTATGAAACATGCGCCAAATAGTATCAAAACGATGTA 1480
Db 1363 GCCATGTCATCCAGTGTCTTCTATGAAACATGCGCCAAATAGTATCAAAACGATGTA 1422
Qy 1481 AATTATCAATTTACCAAAATGTCTGTAACCGAGTGAATGACAGATGATGAT 1540
Db 1423 AATTATCAATTTACCAAAATGTCTGTAACCGAGTGAATGACAGATGATGAT 1482
Qy 1541 GTTATGTTATTCGAAACAGATGTTGGACCGTTCTTAAAGTATGTTCAATTCCTAAGAG 1600
Db 1483 GTTATGTTATTCGAAACAGATGTTGGACCGTTCTTAAAGTATGTTCAATTCCTAAGAG 1542
Qy 1601 ACTTGTA 1608
Db 1543 ACTTGTA 1550

RESULT 15
US-09-652-918-7516
; Sequence 7516, Application US/09652918
; GENERAL INFORMATION:
; APPLICANT: Galvin, Katherine
; APPLICANT: Holzman, Douglas A.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; FILE REFERENCE: 1600.1187-001
; CURRENT APPLICATION NUMBER: US/09/652,918
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,130
; PRIOR FILING DATE: 1999-08-30
; NUMBER OF SEQ ID NOS: 8985
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7516
; LENGTH: 2256
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(2256)
; OTHER INFORMATION: n = A,T,C or G
US-09-652-918-7516

Query Match 51.2%; Score 1388; DB 28; Length 2256;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1388; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 221 GTGTGCTTTTCTGCGGAGATATTACTTACAGCAAGAGCAAACTATCAGATGGAGAAC 280
Db 153 GTGTGCTTTTCTGCGGAGATATTACTTACAGCAAGAGCAAACTATCAGATGGAGAAC 222
Qy 281 AATGTCGAAGCTGAATTTATCCACGAAGAAATGTTGGAATCCAAATGATGCT 340
Db 223 AATGTCGAAGCTGAATTTATCTACGAAGAAATGTTGGAATCCAAATGATGCT 282
Qy 341 TTCAATGCTTGGCCAAAGCTCCAGTATATACCTTCTTTGGATGAGGAACGAGT 400
Db 283 TTCAATGCTTGGCCAAAGCTCCAGTATATACCTTCTTTGGATGAGGAACGAGT 342
Qy 401 AGGCTGATGTTGAGCAAGAGATCAATATTTTCATTCGACCTGGTTAATATCAAGAT 460
Db 343 AGGCTGATGTTGAGCAAGAGATCAATATTTTCATTCGACCTGGTTAATATCAAGAT 402
Qy 461 TTTCAAAAGATTTGTGTGGCCAGTATCTTACACCGAAGATGATGCAAGTGGCTGGA 520
Db 403 TTTCAAAAGATTTGTGTGGCCAGTATCTTACACCGAAGATGATGCAAGTGGCTGGA 462
Qy 521 AAAGCATCTGAAAGATGTGCTAATTTTCATCAAGGTACTTAAAGGCATATATCAGACT 580
Db 463 AAAGCATCTGAAAGATGTGCTAATTTTCATCAAGGTACTTAAAGGCATATATCAGACT 522
Qy 581 CACTTGAACGCTGTGGAACGGGGCTTTTCATCCAAATTTGCACTTACATTGAATTGGA 640
Db 523 CACTTGAACGCTGTGGAACGGGGCTTTTCATCCAAATTTGCACTTACATTGAATTGGA 582
Qy 641 CATCATCTGAGACAAATTTTAACTGAGGAATCAATTTTGAAGAAAGGCGCTGGG 700
Db 583 CATCATCTGAGACAAATTTTAACTGAGGAATCAATTTTGAAGAAAGGCGCTGGG 642
Qy 701 AAGAGTCATATGACCCCTAGCTGTGACAGATCCCTTTAATATGATGAGAAATTAAC 760
Db 643 AAGAGTCATATGACCCCTAGCTGTGACAGATCCCTTTAATATGATGAGAAATTAAC 702
Qy 761 TCTGGAACCTGACGATTTTATGGGGGAGACTTGTGATCTTCCGAATCTTGGGAC 820
Db 703 TCTGGAACCTGACGATTTTATGGGGGAGACTTGTGATCTTCCGAATCTTGGGAC 762
Qy 821 CACCAACCAATCAGACAGACAGATGATTCAGGTGGCTCATGATCCAAAGTTCAAT 880
Db 763 CACCAACCAATCAGACAGACAGATGATTCAGGTGGCTCATGATCCAAAGTTCAAT 822
Qy 881 AGTGCCCACTCATCTCAGAGATGACATCTGAAAGTGAACAAAGTACTTTTCTTC 940
Db 823 AGTGCCCACTCATCTCAGAGATGACATCTGAAAGTGAACAAAGTACTTTTCTTC 882
Qy 941 CGTGAATAATGCAATAGATGAGAAACATCTGAAAGTGAACAAAGTACTTTTCTTC 1000
Db 883 CGTGAATAATGCAATAGATGAGAAACATCTGAAAGTGAACAAAGTACTTTTCTTC 942
Qy 1001 ATATGCAAGATGACTTTGAGGGGACAGAAAGTCTGGAATTAATGGAACAATTCCTC 1060
Db 943 ATATGCAAGATGACTTTGAGGGGACAGAAAGTCTGGAATTAATGGAACAATTCCTC 1002
Qy 1061 AAAGCTGCTGATTTGCTCAGTGCAGAGTCCAAATGGCATGACACTATTTGATGA 1120
Db 1003 AAAGCTGCTGATTTGCTCAGTGCAGAGTCCAAATGGCATGACACTATTTGATGA 1062
Qy 1121 CTGCAAGATGATTTCTAATGAACTTTAAAGATCTTAAATTCAGTTGATATGAGAGT 1180
Db 1063 CTGCAAGATGATTTCTAATGAACTTTAAAGATCTTAAATTCAGTTGATATGAGAGT 1122
Qy 1181 TTTAGACTTCCAGTAACATTTTCAAGGATCAGCCGTGTATGATATGATGATGAT 1240
Db 1123 TTTAGACTTCCAGTAACATTTTCAAGGATCAGCCGTGTATGATATGATGATGAT 1182
Qy 1241 GTGAGAGGGGTCTCTGTGTCATATGCCCCAGGATGACCCAACTATCAATGGGGT 1300
Db 1183 GTGAGAGGGGTCTCTGTGTCATATGCCCCAGGATGACCCAACTATCAATGGGGT 1242
Qy 1301 CCTATCAAGAAAGTCCCTATCCAGCCAGGAACTTGTCCAGCAAAACATTTGGT 1360

Db	1243	CCTTATCAAGGAAGAGTCCCTCATCCAGGCCAAGAACTGTGCCAGCAAAACATTGGT	13020
QY	1361	GGTTTGGACTCTACAAAGGACCTTCCCTGATGATGTTATPACCTTTGCAGAAATCATCCA	14200
Db	1303	GGTTTGGACTCTACAAAGGACCTTCCCTGATGATGTTATPACCTTTGCAGAAAGTATATCA	13620
QY	1421	GCCATGTGCAATCCGATGGTTCCTPAGAAACAATGGCCCAATAGTGAATCAAAAACGGAATGA	14800
Db	1363	GCCATGTGCAATCCGATGGTTCCTPAGAAACAATGGCCCAATAGTGAATCAAAAACGGAATGA	14220
QY	1481	AATTATCAATTTCACAAATGTCCTGAGACCGAGTGTGAGAGATGAGACAGTATGAT	15400
Db	1423	AATTATCAATTTCACAAATGTCCTGAGACCGAGTGTGAGAGATGAGACAGTATGAT	14820
QY	1561	GTTATGTTTATCGGACACAGATGTTGGACCGTTCTTAAAGTATGTTCAATTCTTAAGAG	16000
Db	1483	GTTATGTTTATCGGACACAGATGTTGGACCGTTCTTAAAGTATGTTCAATTCTTAAGAG	15420
QY	1601	ACTTGGA 1608	
Db	1543	ACTTGGA 1550	

```

RESULT 16
US-09-724-676-18078
; Sequence 18078, Application US/09724676
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 129181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18078
; LENGTH: 944
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676-18078

```

Query Match	Similarity	Score	ID#	Length
Beet Local	100.0%	100.0%	0	0
Matches	817	Conservative	0	Mismatches
			0	Indels
			0	Gaps
			0	
QY	193	CTGCAGCATGGGCTGGTTAACTAGGATGTGTCTTTCTGGGAGATTACTTAACGC	252	
Db	9	CTGCAGCATGGGCTGGTTAACTAGGATGTGTCTTTCTGGGAGATTACTTAACGC	68	
QY	253	AAGGCAAACTATCAGAAATGGGAGAACAAATGTGCCAAGGCTGAATATCTTCCAAAGA	312	
Db	69	AAGGCAAACTATCAGAAATGGGAGAACAAATGTGCCAAGGCTGAATATCTTCCAAAGA	128	
QY	313	AATGTGGAAATCCAAACATGTGATCACTTTCATGGCTTGGCCAAACAGCTCAGTTATCA	372	
Db	129	AATGTGGAAATCCAAACATGTGATCACTTTCATGGCTTGGCCAAACAGCTCAGTTATCA	188	
QY	373	TACCTTCTTTTGATGAGAAACGAGTAGCGCTGATGTTGGAGCAAAAGATCAATATT	432	
Db	189	TACCTTCTTTTGATGAGAAACGAGTAGCGCTGATGTTGGAGCAAAAGATCAATATT	248	
QY	433	TTCAATGCACTGGGTTATATCAAGGATTTTCAAAGATTTGTGGCCAGATCTTACAC	492	
Db	249	TTCAATGCACTGGGTTATATCAAGGATTTTCAAAGATTTGTGGCCAGATCTTACAC	308	
QY	493	CAGAAGAGATGAATGCAAGTGCGCTGGAAAAGACATCTGAAAGATGTGCTAATTTTCA	552	
Db	309	CAGAAGAGATGAATGCAAGTGCGCTGGAAAAGACATCTGAAAGATGTGCTAATTTTCA	368	
QY	553	CAAGGTCTTAAGGATTAATTCAGACCTCACTTGAAGCCCTGTGAAACGGGGGCTTTTCA	612	
Db	369	CAAGGTCTTAAGGATTAATTCAGACCTCACTTGAAGCCCTGTGAAACGGGGGCTTTTCA	428	

QY	61.3	TCGATATTTGCACCTACATTTGAATTTGGACATCATCTCTGGAGCAATATATTTTAAAGCTGG	672
Db	429	TCGATATTTGCACCTACATTTGAATTTGGACATCATCTCTGGAGCAATATATTTTAAAGCTGG	488
QY	673	GAACCTACATTTTGAACCGGCGCTGGAGAGAGTCCATATGACCCCTTAAGCTGTGACAG	732
Db	489	GAACCTACATTTTGAACCGGCGCTGGAGAGAGTCCATATGACCCCTTAAGCTGTGACAG	548
QY	723	ATCCCTTTTAAATAGATGGAGAAATTATCTCTGGAACTGCACTGATTTTATGGGGGAGAG	792
Db	549	ATCCCTTTTAAATAGATGGAGAAATTATCTCTGGAACTGCACTGATTTTATGGGGGAGAG	608
QY	793	CTTTGGCTATCTTCGAACTCTTGGGCAACCAACCCAAATCAGAGACAGACAGATGATTC	852
Db	609	CTTTGGCTATCTTCGAACTCTTGGGCAACCAACCCAAATCAGAGACAGACAGATGATTC	668
QY	853	CAGGTGGCTCATGATCCAAAGTTGATTAGTGCCCAACCTCATCTCAGAGAGTGACAAATCC	912
Db	669	CAGGTGGCTCATGATCCAAAGTTGATTAGTGCCCAACCTCATCTCAGAGAGTGACAAATCC	728
QY	913	TGAAGATGACAAAGTATATCTTTTCTTCGTCGAAAAATGCATATGATGAGAAACACTCTGG	972
Db	729	TGAAGATGACAAAGTATATCTTTTCTTCGTCGAAAAATGCATATGATGAGAAACACTCTGG	788
QY	973	AAAAGCTACTACGCTAGATGATGCTCAGATATGCAAG	1009
Db	789	AAAAGCTACTACGCTAGATGATGCTCAGATATGCAAG	825

```

RESULT 17
US-09-724-676A-18078
; Sequence 18078, Application US/09724676A
; GENERAL INFORMATION:
; APPLICANT: CompuGen LTD
; TITLE OF INVENTION: Variants of alternative splicing
; FILE REFERENCE: 123181.4 CompuGen
; CURRENT APPLICATION NUMBER: US/09/724,676A
; CURRENT FILING DATE: 2000-11-28
; NUMBER OF SEQ ID NOS: 97222
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18078
; LENGTH: 944
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-724-676A-18078

```

Query Match	Similarity	Score	DB	Length	Matches	Conservative	Mismatches	Indels	Gaps
30.2%	100.0%	817	32	944	0	0	0	0	0
Best Local									
Match									
817									
QY	193	CTGCAGCATGGGCTGGTTAACTAGGATTTGCTCTTTTCTGGGAGTTACTTACACG	252						
DB	9	CTGCAGCATGGGCTGGTTAACTAGGATTTGCTCTTTTCTGGGAGTTACTTACACG	68						
QY	253	AAGAGCAACTATCAGAAATGGGAGAACAAATGTGCGCAAGCTGAATTATCTTACAAAGA	312						
DB	69	AAGAGCAACTATCAGAAATGGGAGAACAAATGTGCGCAAGCTGAATTATCTTACAAAGA	128						
QY	313	AATGTGGAAATCCAAACAATGTGATCACTTTTCAATGGCTTGGCCACAGCTCCAGTTATCA	372						
DB	129	AATGTGGAAATCCAAACAATGTGATCACTTTTCAATGGCTTGGCCACAGCTCCAGTTATCA	188						
QY	373	TACCTTCTTTTGGATGAGAACGGAGTAGGCTGTATGTGTGGAGCAAAAGATCAACATATT	432						
DB	189	TACCTTCTTTTGGATGAGAACGGAGTAGGCTGTATGTGTGGAGCAAAAGATCAACATATT	248						
QY	433	TTCAATTCGACCTGGTTAAATATCAAGGATTTTCAAAAAGATGTGTGAGCCAGTATCTTACAC	492						
DB	249	TTCAATTCGACCTGGTTAAATATCAAGGATTTTCAAAAAGATGTGTGAGCCAGTATCTTACAC	308						
QY	493	CAGAAAGATGAAATGCAAGTGGGCTGGAAAAAGACATCTCTGAAGAAGATGTGCTAATTTTCA	552						
DB	309	CAGAAAGATGAAATGCAAGTGGGCTGGAAAAAGACATCTCTGAAGAAGATGTGCTAATTTTCA	368						

QY 553 CAAGTACTTAAAGCATATATATGAGTCACTTGTAGCGCTGTGAAACGGGGCTTTTGA 612
DB 369 CAAGTACTTAAAGCATATATATGAGTCACTTGTAGCGCTGTGAAACGGGGCTTTTGA 428
QY 613 TCCAAATTTGACCTTAAATTTGAATTTGAGATCTCTGAGGACAAATTTTAAAGCTGGA 672
DB 429 TCCAAATTTGACCTTAAATTTGAATTTGAGATCTCTGAGGACAAATTTTAAAGCTGGA 488
QY 673 GAATCACTATTTTGAAGGCGCGTGGAGAGTCCATATGACCTTAAAGCTGTGACAGC 732
DB 489 GAATCACTATTTTGAAGGCGCGTGGAGAGTCCATATGACCTTAAAGCTGTGACAGC 548
QY 733 ATCCCTTTTAAATGATGAGAAATTTATCTGTGAACTGTGAGTGTATTTTAAAGCTGGA 792
DB 549 ATCCCTTTTAAATGATGAGAAATTTATCTGTGAACTGTGAGTGTATTTTAAAGCTGGA 608
QY 793 CTTCCTTATCTTCCGAATCTTTGGGACACCAACCAATCAGACAGACAGATGATTC 852
DB 609 CTTCCTTATCTTCCGAATCTTTGGGACACCAACCAATCAGACAGACAGATGATTC 668
QY 853 CAGGTGGCTCAATGATCCAAAGTTCATTTAGTCCCACTCATCTCAGAGAGTGAACAATCC 912
DB 669 CAGGTGGCTCAATGATCCAAAGTTCATTTAGTCCCACTCATCTCAGAGAGTGAACAATCC 728
QY 913 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAAAATGCAATGATGAGAACTGTGG 972
DB 729 TGAAGATGACAAAGTATATCTTTTCTCCGTGAAAAATGCAATGATGAGAACTGTGG 788
QY 973 AAAAGTACTCAGCTGTAATATGTCATATGTCAG 1009
DB 789 AAAAGTACTCAGCTGTAATATGTCATATGTCAG 825

RESULT 18

US-09-864-761-30992
; Sequence 30992, Application US/09864761
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME- DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30

;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annonex Sequence Listing Engine vers. 1.1
;; SEQ ID NO 30992
;; LENGTH: 456
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; PEATRE:
;; OTHER INFORMATION: MAP TO AC006322.2
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.73
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.55
;; OTHER INFORMATION: EST HUMAN HIT: BF700780.1, EVALUE 0.00e+00
;; OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 2.00e-88
;; OTHER INFORMATION: NT HIT: g111421514, EVALUE 0.00e+00
US-09-864-761-30992

Query Match 15.0%; Score 405; DB 36; Length 456;
Best Local Similarity 99.8%; Pred. No. 4,6e-172;

Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2060 ATCAGATGATGATCATATATCATGAGACAGATCAAGGCTTTGCTAGTATGATCA 2119
DB 1 ATCAGATGATGATCATATATCATGAGACAGATCAAGGCTTTGCTAGTATGATCA 60
QY 2120 CAGAGGATTCAGGCAATTAATCTCTGCGATCGGTGGAACATGGTTATATCAACTCTT 2179
DB 61 CAGAGGATTCAGGCAATTAATCTCTGCGATCGGTGGAACATGGTTATATCAACTCTT 120
QY 2180 CTAAAGTAACTCCCTGGAAGTATGACACAGACAGATTTGGAAGAACTTTTATTAAGT 2239
DB 121 CTAAAGTAACTCCCTGGAAGTATGACACAGACAGATTTGGAAGAACTTTTATTAAGT 180
QY 2240 GATGATGAGATGCTCTTAAGACCAAGAAATGTCATTAAGATGACATGACCAAG 2299
DB 181 GATGATGAGATGCTCTTAAGACCAAGAAATGTCATTAAGATGACATGACCAAG 240
QY 2300 GTCTGTACAGAGCTTATGAGAGTCTCAACCAACCCATCTTAACAGATGATGAG 2359
DB 241 GTCTGTACAGAGCTTATGAGAGTCTCAACCAACCCATCTTAACAGATGATGAG 300
QY 2360 TTCTGTGAACAAGTTTGAAGGAGCCGAAACCAAGCTGGGCAAGGCGAGACATACC 2419
DB 301 TTCTGTGAACAAGTTTGAAGGAGCCGAAACCAAGCTGGGCAAGGCGAGACATACC 360
QY 2420 CCAGGAAACAGTAACTTAAGAGCACTTACAGAAATTAAGAAAGGTAAGAAACAGAGG 2479
DB 361 CCAGGAAACAGTAACTTAAGAGCACTTACAGAAATTAAGAAAGGTAAGAAACAGAGG 420
QY 2480 ACCCAAGATTTGAGAGGCAACCCAGAGAGTCTTGA 2515
DB 421 ACCCAAGATTTGAGAGGCAACCCAGAGAGTCTTGA 456

RESULT 19

US-10-203-135-19786
; Sequence 19786, Application US/10203135
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME- DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: ANALYSIS OF GENE EXPRESSION IN HUMAN LUNG

```
FILE REFERENCE: PB 0004 WO 5
CURRENT FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 37012
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 19786
LENGTH: 456
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006322.2
FEATURE:
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.73
FEATURE:
OTHER INFORMATION: EST_HUMAN HIT: BF700780.1, EVALUE 0.00e+00
FEATURE:
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 2.00e-88
FEATURE:
OTHER INFORMATION: NT HIT: g111421514, EVALUE 0.00e+00
US-10-203-135-19786
```

```
Query Match 15.0%; Score 405; DB 48; Length 456;
Best Local Similarity 99.8%; Pred. No. 4.6e-172;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2060 ATCAGAGTGATGATCATATCATCAGACAGATCAAGGCTTGTAGTACTTACAA 2119
DB 1 ATCAGAGTGATGATCATATCATCAGACAGATCAAGGCTTGTAGTACTTACAA 60
QY 2120 CAGAAGATTGAGCAATTAACCTCTGCGATGCGGTGGAACATGGTTCTTACAACTCTT 2179
DB 61 CAGAAGATTGAGCAATTAACCTCTGCGATGCGGTGGAACATGGTTCTTACAACTCTT 120
QY 2180 CTTAAGTAACTCTGGAAGTCAATGACACAGAGCATTTGGAAGAACTTCTTCAATAAGAT 2239
DB 121 CTTAAGTAACTCTGGAAGTCAATGACACAGAGCATTTGGAAGAACTTCTTCAATAAGAT 180
QY 2240 GATGATGAGATGGCTCTTAAGACCAAGAAATGTCATATAGTACACCTTACCCGAGAG 2299
DB 181 GATGATGAGATGGCTCTTAAGACCAAGAAATGTCATATAGTACACCTTACCCGAGAG 240
QY 2300 GTCTGTACAGAGACTTCAATGAGCTCATCAACCCCAATCTCAACAGATGATGAG 2359
DB 241 GTCTGTACAGAGACTTCAATGAGCTCATCAACCCCAATCTCAACAGATGATGAG 300
QY 2360 TTCTGTGAACAAGTTTGAAGAAAGGACCGAACAACGTCGGCAAAAGCCAGAGACATACC 2419
DB 301 TTCTGTGAACAAGTTTGAAGAAAGGACCGAACAACGTCGGCAAAAGCCAGAGACATACC 360
QY 2420 CCAGGAGACATTAACAATGAAAGCCTTACAAAGAAATTAAGAAAGTGAAGACGAGAG 2479
DB 361 CCAGGAGACATTAACAATGAAAGCCTTACAAAGAAATTAAGAAAGTGAAGACGAGAG 420
QY 2480 ACCCAAGATTTGAGAGGGCCACCGAGAGTGTCTGA 2515
DB 421 ACCCAAGATTTGAGAGGGCCACCGAGAGTGTCTGA 456
```

RESULT 20

```
US-10-203-136-20203
Sequence 20203, Application US/10203136
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER
FILE REFERENCE: PB 0004 WO 3
CURRENT APPLICATION NUMBER: US/10/203,136
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38578
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 20203
LENGTH: 456
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006322.2
FEATURE:
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.55
FEATURE:
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.73
FEATURE:
OTHER INFORMATION: EST_HUMAN HIT: BF700780.1, EVALUE 0.00e+00
FEATURE:
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 2.00e-88
FEATURE:
OTHER INFORMATION: NT HIT: g111421514, EVALUE 0.00e+00
US-10-203-136-20203
```

```
Query Match 15.0%; Score 405; DB 48; Length 456;
Best Local Similarity 99.8%; Pred. No. 4.6e-172;
Matches 455; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2060 ATCAGAGTGATGATCATATCATCAGACAGATCAAGGCTTGTAGTACTTACAA 2119
DB 1 ATCAGAGTGATGATCATATCATCAGACAGATCAAGGCTTGTAGTACTTACAA 60
QY 2120 CAGAAGATTGAGCAATTAACCTCTGCGATGCGGTGGAACATGGTTCTTACAACTCTT 2179
DB 61 CAGAAGATTGAGCAATTAACCTCTGCGATGCGGTGGAACATGGTTCTTACAACTCTT 120
QY 2180 CTTAAGTAACTCTGGAAGTCAATGACACAGAGCATTTGGAAGAACTTCTTCAATAAGAT 2239
DB 121 CTTAAGTAACTCTGGAAGTCAATGACACAGAGCATTTGGAAGAACTTCTTCAATAAGAT 180
QY 2240 GATGATGAGATGGCTCTTAAGACCAAGAAATGTCATATAGTACACCTTACCCGAGAG 2299
DB 181 GATGATGAGATGGCTCTTAAGACCAAGAAATGTCATATAGTACACCTTACCCGAGAG 240
QY 2300 GTCTGTACAGAGACTTCAATGAGCTCATCAACCCCAATCTCAACAGATGATGAG 2359
DB 241 GTCTGTACAGAGACTTCAATGAGCTCATCAACCCCAATCTCAACAGATGATGAG 300
QY 2360 TTCTGTGAACAAGTTTGAAGAAAGGACCGAACAACGTCGGCAAAAGCCAGAGACATACC 2419
DB 301 TTCTGTGAACAAGTTTGAAGAAAGGACCGAACAACGTCGGCAAAAGCCAGAGACATACC 360
```



```

1  APPLICANT: Randhwa, Gurpreet
2  APPLICANT: Sidhu, Navjivan
3  APPLICANT: Smith, Benjamin
4  APPLICANT: Smythe, Ashleigh
5  APPLICANT: Trach, Joe
6  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
7  TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
8  FILE REFERENCE: 740CIP
9  CURRENT APPLICATION NUMBER: US/09/733, 811
10 CURRENT FILING DATE: 2000-12-08
11 PRIOR APPLICATION NUMBER: 09/321, 214
12 PRIOR FILING DATE: 1999-05-26
13 PRIOR APPLICATION NUMBER: 60/088, 041
14 PRIOR FILING DATE: 1998-06-02
15 NUMBER OF SEQ ID NOS: 31906
16 SOFTWARE: FastSeq for Windows Version 3.0
17 SEQ ID NO 10943
18 LENGTH: 379
19 TYPE: DNA
20 ORGANISM: Homo sapiens
21 US-09-733-811-10943

```

	Query Match	Similarity	100.0%	Score 379;	DB 32;	Length 379;
	Best Local	Similarity	100.0%	Pred. No.	2.8e-160;	
	Matches	379;	Conservative	0;	Mismatches	0; Indels
Oy	2196	AAGTCATTGACACAGAGCATTTGGAGAACACTTCTTCATTAAGATGATGAAGATGGCT				2255
Dd	1	AAGTCATTGACACAGAGCATTTGGAGAACACTTCTTCATTAAGATGATGAAGATGGCT				60
Oy	2256	CTAAGACCMAAGAAATGTCCCAATAGCANTGACACCTTAGGCCAGAAGGTCGTGCACAGACT				2315
Dd	61	CTAAGACCMAAGAAATGTCCCAATAGCANTGACACCTTAGGCCAGAAGGTCGTGCACAGACT				120
Oy	2316	TCAATGACGCTCATCAACCAACCCCAAATCTCAACAGATGATGAAGTTCGTGACAAAGTT				2375
Dd	121	TCAATGACGCTCATCAACCAACCCCAAATCTCAACAGATGATGAAGTTCGTGACAAAGTT				180
Oy	2376	GAAAAAGGAGCGAAAACMAGTGCGGCAAAAGGCCAGACATATCCCAGAGGAACAGTAACA				2435
Dd	181	GAAAAAGGAGCGAAAACMAGTGCGGCAAAAGGCCAGACATATCCCAGAGGAACAGTAACA				240
Oy	2436	AATGAGACCTTACATAAATAAATAAGGTAGAAACAGAGAGCCACAGATTGAGA				2495
Dd	241	AATGAGACCTTACATAAATAAATAAGGTAGAAACAGAGAGCCACAGATTGAGA				300
Oy	2496	GGGACCCCCAGAGAGTCTCTGAGCTGATTAACCTCTAGAAACCTCAAAACAGTAGAAACTTG				2555
Dd	301	GGGACCCCCAGAGAGTCTCTGAGCTGATTAACCTCTAGAAACCTCAAAACAGTAGAAACTTG				360
Oy	2556	CCTAGACACATAACTGGAAA	2574			
Dd	361	CCTAGACACATAACTGGAAA	379			
RESULT 26						
US-09-733-811A-10943						
Sequence 10943, Application US/09733811A						
GENERAL INFORMATION:						
APPLICANT: Hyseq, Inc.						
APPLICANT: Drmanac, Radoje						
APPLICANT: Cirvenjakov, Radomir						
APPLICANT: Dickson, Mark						
APPLICANT: Drmanac, Snezana						
APPLICANT: Labat, Ivan						
APPLICANT: Leshkowitz, Dena						
APPLICANT: Kita, David						
APPLICANT: Ford, John						
TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED						
TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES						
FILE REFERENCE: 20411-740						
CURRENT APPLICATION NUMBER: US/09/733.811A						
CURRENT FILING DATE: 2000-12-08						

```

? PRIOR APPLICATION NUMBER: 09/321,214
? PRIOR FILING DATE: 1999-05-26
? NUMBER OF SEQ. ID NOS: 31906
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ. ID NO. 10943
? LENGTH: 379
? TYPE: DNA
? ORGANISM: Homo sapiens
? OS-09-73-811A-10943

```

```

Query Match      14.0%; Score 379; DB 37; Length 379;
Best Local Similarity 100.0%; Pred. No. 2,8e-160;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2196 AAGTCATTGACACAGACCAATTTGGAGAAGCTTCTTCATTAAGATGATGATGAGATGAGCT 2255
Db      1  AAGTCATTGACACAGAGCATTTGGAGAAGCTTCTTCATTAAGATGATGATGAGATGAGCT 60

QY      2256 CTAAAGACCCAAAGAAATGTCCCAATGCGATGACACTTACGCCAGAAAGTCTGGTACAGAGCT 2315
Db      61  CTAAAGACCCAAAGAAATGTCCCAATGCGATGACACTTACGCCAGAAAGTCTGGTACAGAGCT 120

QY      2316 TCATGACAGCTCATAAACACCCCAATCTCAACACAGATGAGTGAATTCGTGTAACAAGTT 2375
Db      121  TCATGACAGCTCATAAACACCCCAATCTCAACACAGATGAGTGAATTCGTGTAACAAGTT 180

QY      2376 GGAAGAAAGGACCCGAAACCAACGTCGGCAAAAGCCAGACATACCCCGAGGAACAGTAAACA 2435
Db      181  GGAAGAAAGGACCCGAAACCAACGTCGGCAAAAGCCAGACATACCCCGAGGAACAGTAAACA 240

QY      2436 AATGGAAGCATCTTACAGAAATAAGAAAGGTAGAAAACAGAGAGACCCAGAAATTTGAGA 2495
Db      241  AATGGAAGCATCTTACAGAAATAAGAAAGGTAGAAAACAGAGAGACCCAGAAATTTGAGA 300

QY      2496 GGGGACCCAGAGAGTGTCTGAGCTGCATTAACCTCTAGAAAACCTCAACAAAGTAGAAAACCTG 2555
Db      301  GGGGACCCAGAGAGTGTCTGAGCTGCATTAACCTCTAGAAAACCTCAACAAAGTAGAAAACCTG 360

QY      2556 CCTAGACAAATTAAGTGTGAAA 2574
Db      361  CCTAGACAAATTAAGTGTGAAA 379

RESULT 27
US-09-898-45112
; Sequence 45112; Application US/09898888
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS cDNA
; FILE REFERENCE: 20411-748CON1
; CURRENT APPLICATION NUMBER: US/09/898,888
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 09/340,623
; PRIOR FILING DATE: 1999-06-28
; NUMBER OF SEQ ID NOS: 45207
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 45112
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-898-888-45112

Query Match      14.0%; Score 379; DB 37; Length 379;
Best Local Similarity 100.0%; Pred. No. 2.8e-160;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2196 AAGTCATTGACACAGACCAATTTGGAGAAGCTTCTTCATTAAGATGATGATGAGATGAGCT 2255
Db      1  AAGTCATTGACACAGAGCATTTGGAGAAGCTTCTTCATTAAGATGATGATGAGATGAGCT 60

QY      2256 CTAAAGACCCAAAGAAATGTCCCAATGCGATGACACTTACGCCAGAAAGTCTGGTACAGAGCT 2315
Db      61  CTAAAGACCCAAAGAAATGTCCCAATGCGATGACACTTACGCCAGAAAGTCTGGTACAGAGCT 120

QY      2316 TCATGACAGCTCATAAACACCCCAATCTCAACACAGATGAGTGAATTCGTGTAACAAGTT 2375
Db      121  TCATGACAGCTCATAAACACCCCAATCTCAACACAGATGAGTGAATTCGTGTAACAAGTT 180

QY      2376 GGAAGAAAGGACCCGAAACCAACGTCGGCAAAAGCCAGACATACCCCGAGGAACAGTAAACA 2435
Db      181  GGAAGAAAGGACCCGAAACCAACGTCGGCAAAAGCCAGACATACCCCGAGGAACAGTAAACA 240

QY      2436 AATGGAAGCATCTTACAGAAATAAGAAAGGTAGAAAACAGAGAGACCCAGAAATTTGAGA 2495
Db      241  AATGGAAGCATCTTACAGAAATAAGAAAGGTAGAAAACAGAGAGACCCAGAAATTTGAGA 300

QY      2496 GGGGACCCAGAGAGTGTCTGAGCTGCATTAACCTCTAGAAAACCTCAACAAAGTAGAAAACCTG 2555
Db      301  GGGGACCCAGAGAGTGTCTGAGCTGCATTAACCTCTAGAAAACCTCAACAAAGTAGAAAACCTG 360

QY      2556 CCTAGACAAATTAAGTGTGAAA 2574
Db      361  CCTAGACAAATTAAGTGTGAAA 379

```

```
Db      61 CTAAGACCAAAAGATGTCCATAGCATGACACCTTACCGCAAGAGTCTGTACAGAGACT 120
QY      2316 TCATGAGCTCATCAACCAACCCCAATCTCAACAGATGATGATTTGAACTTT 2375
Db      121 TCATGAGCTCATCAACCAACCCCAATCTCAACAGATGATGATTTGAACTTT 180
QY      2376 GGAAGAGGAGCCGAAACCAACGTCGCAAGGCCAGAGCATTCACCCAGGAGCACTTAC 2435
Db      181 GGAAGAGGAGCCGAAACCAACGTCGCAAGGCCAGAGCATTCACCCAGGAGCACTTAC 240
QY      2436 AATGGAAGCACTTACAAAGAAATTAAGAAAGGTAGAAACAGAGAGACCCAGATTTGAG 2495
Db      241 AATGGAAGCACTTACAAAGAAATTAAGAAAGGTAGAAACAGAGAGACCCAGATTTGAG 300
QY      2496 GGGCACCACCGAGAGTGTCTGAGCTGATTAACCTTACAAACCTCAAAAGATAGAACTTG 2555
Db      301 GGGCACCACCGAGAGTGTCTGAGCTGATTAACCTTACAAACCTCAAAAGATAGAACTTG 360
QY      2556 CCTAGACAAATACTGGAAA 2574
Db      361 CCTAGACAAATACTGGAAA 379
```

```
RESULT 28
US-09-898-888A-45112
; Sequence 45112, Application US/09898888A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE REFERENCE: 20411-748CON1
; CURRENT APPLICATION NUMBER: US/09/898, 888A
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/340,623
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/205,070
; NUMBER OF SEQ ID NOS: 45207
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 45112
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-898-888A-45112
```

```
Query Match      14.0%; Score 379; DB 37; Length 379;
Best Local Similarity 100.0%; Pred. No. 2.8e-160;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2196 AAGTCATTGACACAGAGATTGGAAGAACTTCTCATTAAGATGATGATGATGCT 2255
Db      1 AAGTCATTGACACAGAGATTGGAAGAACTTCTCATTAAGATGATGATGATGCT 60
QY      2256 CTAAGACCAAAAGATGTCCATAGCATGACACCTTACCGCAAGAGTCTGTACAGAGACT 2315
Db      61 CTAAGACCAAAAGATGTCCATAGCATGACACCTTACCGCAAGAGTCTGTACAGAGACT 120
QY      2316 TCATGAGCTCATCAACCAACCCCAATCTCAACAGATGATGATTTGAACTTT 2375
Db      121 TCATGAGCTCATCAACCAACCCCAATCTCAACAGATGATGATTTGAACTTT 180
QY      2376 GGAAGAGGAGCCGAAACCAACGTCGCAAGGCCAGAGCATTCACCCAGGAGCACTTAC 2435
Db      181 GGAAGAGGAGCCGAAACCAACGTCGCAAGGCCAGAGCATTCACCCAGGAGCACTTAC 240
QY      2436 AATGGAAGCACTTACAAAGAAATTAAGAAAGGTAGAAACAGAGAGACCCAGATTTGAG 2495
Db      241 AATGGAAGCACTTACAAAGAAATTAAGAAAGGTAGAAACAGAGAGACCCAGATTTGAG 300
QY      2496 GGGCACCACCGAGAGTGTCTGAGCTGATTAACCTTACAAACCTCAAAAGATAGAACTTG 2555
Db      301 GGGCACCACCGAGAGTGTCTGAGCTGATTAACCTTACAAACCTCAAAAGATAGAACTTG 360
```

```
QY      2556 CCTAGACAAATACTGGAAA 2574
Db      361 CCTAGACAAATACTGGAAA 379
```

```
RESULT 29
US-09-975-640-10943
; Sequence 10943, Application US/09975640
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jessen, Aaron
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroya, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nelken, Sarah
; APPLICANT: Nguyen, Kody
; APPLICANT: Nguyen, Lynne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sidhu, Navyian
; APPLICANT: Smith, Benjamin
; APPLICANT: Smyche, Ashleigh
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Wachter, Adam
; APPLICANT: Wu, James
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 740CIP
; CURRENT APPLICATION NUMBER: US/09/975,640
; PRIOR FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 09/733,811
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/088,041
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10943
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-975-640-10943
```

```
Query Match      14.0%; Score 379; DB 42; Length 379;
Best Local Similarity 100.0%; Pred. No. 2.8e-160;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2196 AAGTCATTGACACAGAGATTGGAAGAACTTCTCATTAAGATGATGATGATGCT 2255
Db      1 AAGTCATTGACACAGAGATTGGAAGAACTTCTCATTAAGATGATGATGATGCT 60
QY      2256 CTAAGACCAAAAGATGTCCATAGCATGACACCTTACCGCAAGAGTCTGTACAGAGACT 2315
Db      61 CTAAGACCAAAAGATGTCCATAGCATGACACCTTACCGCAAGAGTCTGTACAGAGACT 120
```

```
QY 2316 TCATGACGCTCATCAACCAACCCCAATCTCAACACGATGATGATTTCTGTGAACAAGTTT 2375
DB 121 TCATGACGCTCATCAACCAACCCCAATCTCAACACGATGATGATTTCTGTGAACAAGTTT 180
QY 2376 GAAAAAGGACCGAAAAACAGTGTGGCAAAAGCCAGAGATATCCCAAGGAAACAGTAACA 2435
DB 181 GAAAAAGGACCGAAAAACAGTGTGGCAAAAGCCAGAGATATCCCAAGGAAACAGTAACA 240
QY 2436 AATGAGACCTTACAGAAATAAGAAAGTAGAAACAGAGACCCCAAGATTTTGAGA 2495
DB 241 AATGAGACCTTACAGAAATAAGAAAGTAGAAACAGAGACCCCAAGATTTTGAGA 300
QY 2496 GGGACCCGACGAGTGTCTGAGCTGATTAACCTCTAGAAAACCTCAACAGTAGAAACTTG 2555
DB 301 GGGACCCGACGAGTGTCTGAGCTGATTAACCTCTAGAAAACCTCAACAGTAGAAACTTG 360
QY 2556 CCTAGACATTAACCTGGAAA 2574
DB 361 CCTAGACATTAACCTGGAAA 379

RESULT 30
US-09-975-640A-10943
; Sequence 10943, Application US/09975640A
; GENERAL INFORMATION:
; APPLICANT: Arterburn, Matthew
; APPLICANT: Asghari, Vida
; APPLICANT: Damavandi, Simin
; APPLICANT: Dickson, Mark
; APPLICANT: Drake, Jim
; APPLICANT: Drmanac, Radoje
; APPLICANT: Engleman, Carrie
; APPLICANT: Faulkner, Brandy
; APPLICANT: Fox, Melvin
; APPLICANT: Garcia, Veronica
; APPLICANT: Giedt, Gretchen
; APPLICANT: Jessen, Aaron
; APPLICANT: Jomek, Leni
; APPLICANT: Jones, Lee
; APPLICANT: Kita, David
; APPLICANT: Labat, Ivan
; APPLICANT: Laroza, Mimi
; APPLICANT: Lomelli, Michelle
; APPLICANT: Nelken, Sarah
; APPLICANT: Nguyen, Kody
; APPLICANT: Nguyen, Lynne
; APPLICANT: Nguyen, Phuong
; APPLICANT: Nogra, Margie
; APPLICANT: Palencia, Servando
; APPLICANT: Raisi, Fariba
; APPLICANT: Randhwa, Gurpreet
; APPLICANT: Sidhu, Navjivan
; APPLICANT: Smith, Benjamin
; APPLICANT: Smythe, Ashleigh
; APPLICANT: Tkach, Joe
; APPLICANT: Tran, Lien
; APPLICANT: Verna, Ron
; APPLICANT: Wachter, Adam
; APPLICANT: Wu, James
; APPLICANT: Yim, Kenneth
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 740CIP
; CURRENT APPLICATION NUMBER: US/09/975,640A
; CURRENT FILING DATE: 2001-10-11
; PRIOR APPLICATION NUMBER: 09/733,811
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/088,041
; PRIOR FILING DATE: 1998-06-02
; NUMBER OF SEQ ID NOS: 31906
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10943
; LENGTH: 379
```

```
TYPE: DNA
ORGANISM: Homo sapiens
US-09-975-640A-10943

Query Match 14.0%; Score 379; DB 42; Length 379;
Best Local Similarity 100.0%; Pred. No. 2,8e-160;
Matches 379; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2136 AAGTCATTGACACAGAGCATTTGGAGAACTTTCTTCATTAAGATGATGATGAGTGGCT 2255
DB 1 AAGTCATTGACACAGAGCATTTGGAGAACTTTCTTCATTAAGATGATGATGAGTGGCT 60
QY 2256 CTAGACCCAAAGAAATGTCATATGACATGACACCTAGCCAGAAAGTCTGTACAGAGACT 2315
DB 61 CTAGACCCAAAGAAATGTCATATGACATGACACCTAGCCAGAAAGTCTGTACAGAGACT 120
QY 2316 TCATGACGCTCATCAACCAACCCCAATCTCAACACGATGATGATTTCTGTGAACAAGTTT 2375
DB 121 TCATGACGCTCATCAACCAACCCCAATCTCAACACGATGATGATTTCTGTGAACAAGTTT 180
QY 2376 GAAAAAGGACCGAAAAACAGTGTGGCAAAAGCCAGAGCATATCCCAAGGAAACAGTAACA 2435
DB 181 GAAAAAGGACCGAAAAACAGTGTGGCAAAAGCCAGAGCATATCCCAAGGAAACAGTAACA 240
QY 2436 AATGAGACCTTACAGAAATAAGAAAGTAGAAACAGAGACCCCAAGATTTTGAGA 2495
DB 241 AATGAGACCTTACAGAAATAAGAAAGTAGAAACAGAGACCCCAAGATTTTGAGA 300
QY 2496 GGGACCCGACGAGTGTCTGAGCTGATTAACCTCTAGAAAACCTCAACAGTAGAAACTTG 2555
DB 301 GGGACCCGACGAGTGTCTGAGCTGATTAACCTCTAGAAAACCTCAACAGTAGAAACTTG 360
QY 2556 CCTAGACATTAACCTGGAAA 2574
DB 361 CCTAGACATTAACCTGGAAA 379

RESULT 31
US-09-534-856-4
; Sequence 4, Application US/09534856
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING OR REGULATING STRUCTURAL, SECRETED, AND
; FILE REFERENCE: PD-1015 CIP
; CURRENT APPLICATION NUMBER: US/09/534,856
; CURRENT FILING DATE: 2000-03-24
; PRIOR APPLICATION data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 26334
; SOFTWARE: PERL Program
; SEQ ID NO 4
; LENGTH: 497
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Inocyte ID No: hu01053939
; NAME/KEY: unsure
; LOCATION: 422
; OTHER INFORMATION: a, t, c, g, or other
US-09-534-856-4

Query Match 13.7%; Score 370; DB 23; Length 497;
Best Local Similarity 99.8%; Pred. No. 3.3e-156;
Matches 420; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2163 GGTTCATACAAACTCTTCTTAAGTAAACCTGGAAGTCAATGACAGAGCAATTTGGAG 2222
```


Db 241 CAATCTCAACACAGATGAGTGTCTGTAACAAGTTGGAAAAGGACCGAAAACAAG 300
 QY 2398 TCGGCAAAAGCCAGAGCATACCCCGAGGAACAGTAACAATGGAAGCACTTCA 2451
 Db 301 TCGGCAAAAGCCAGAGCATACCCCGAGGAACAGTAACAATGGAAGCACTTCA 354

RESULT 34

US-09-649-811-736
 ; Sequence 736, Application US/09649811

; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jianshun
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Secrist, Heather
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Meagher, Madeleine Joy
 ; APPLICANT: King, Gordon E.
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
 ; FILE REFERENCE: 210121.471C11
 ; CURRENT APPLICATION NUMBER: US/09/649,811
 ; CURRENT FILING DATE: 2000-08-28
 ; NUMBER OF SEQ ID NOS: 1083
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 736
 ; LENGTH: 354
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(354)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-649-811-736

Query Match 11.2%; Score 303; DB 28; Length 354;
 Best Local Similarity 99.7%; Pred. No. 7,6e-126;
 Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2098 CCTTCTGCTAGCTAGTCTTACACAGAGGATTCAGGCAATTACCTTGGCAGCGGTGA 2157
 Db 1 CCTTCTGCTAGCTAGTCTTACACAGAGGATTCAGGCAATTACCTTGGCAGCGGTGA 60
 QY 2158 ACATGGGTTCAATACAACTCTTAAAGTAACCTTGAAAGTCATTGACACAGACATT 2217
 Db 61 ACATGGGTTCAATACAACTCTTAAAGTAACCTTGAAAGTCATTGACACAGACATT 120
 QY 2218 GGAAGAACTTCTTCAATAAAGATGATGAGATGAGTGGCTTAAAGCAAAAGAAATGTCCA 2277
 Db 121 GGAAGAACTTCTTCAATAAAGATGATGAGATGAGTGGCTTAAAGCAAAAGAAATGTCCA 180
 QY 2278 TAGCATGACCTTACCTACCAAGAGTCTGTGACAGACATTCATGACGCTCATACACACC 2337
 Db 181 TAGCATGACCTTACCTACCAAGAGTCTGTGACAGACATTCATGACGCTCATACACACC 240
 QY 2338 CAATCTCAACAGATGATGAGTGTCTGTGAACAAGTTGGAAAAGGACCGAAAACAAG 2397
 Db 241 CAATCTCAACAGATGATGAGTGTCTGTGAACAAGTTGGAAAAGGACCGAAAACAAG 300
 QY 2398 TCGGCAAAAGCCAGAGCATACCCCGAGGAACAGTAACAATGGAAGCACTTCA 2451
 Db 301 TCGGCAAAAGCCAGAGCATACCCCGAGGAACAGTAACAATGGAAGCACTTCA 354

RESULT 35

US-09-833-263-736
 ; Sequence 736, Application US/09833263

; GENERAL INFORMATION:
 ; APPLICANT: Wang, Aljun
 ; APPLICANT: Clapper, Jonathan D.
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Meagher, Madeleine J.
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
 ; FILE REFERENCE: 210121.471C12
 ; CURRENT APPLICATION NUMBER: US/09/833,263
 ; CURRENT FILING DATE: 2001-04-10
 ; NUMBER OF SEQ ID NOS: 1093
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 736
 ; LENGTH: 354
 ; TYPE: DNA
 ; ORGANISM: Homo sapien
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(354)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-833-263-736

Query Match 11.2%; Score 303; DB 35; Length 354;
 Best Local Similarity 99.7%; Pred. No. 7,6e-126;
 Matches 353; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2098 CCTTCTGCTAGCTAGTCTTACACAGAGGATTCAGGCAATTACCTTGGCAGCGGTGA 2157
 Db 1 CCTTCTGCTAGCTAGTCTTACACAGAGGATTCAGGCAATTACCTTGGCAGCGGTGA 60
 QY 2158 ACATGGGTTCAATACAACTCTTAAAGTAACCTTGAAAGTCATTGACACAGACATT 2217
 Db 61 ACATGGGTTCAATACAACTCTTAAAGTAACCTTGAAAGTCATTGACACAGACATT 120
 QY 2218 GGAAGAACTTCTTCAATAAAGATGATGAGATGAGTGGCTTAAAGCAAAAGAAATGTCCA 2277
 Db 121 GGAAGAACTTCTTCAATAAAGATGATGAGATGAGTGGCTTAAAGCAAAAGAAATGTCCA 180
 QY 2278 TAGCATGACCTTACCTACCAAGAGTCTGTGACAGACATTCATGACGCTCATACACACC 2337
 Db 181 TAGCATGACCTTACCTACCAAGAGTCTGTGACAGACATTCATGACGCTCATACACACC 240
 QY 2338 CAATCTCAACAGATGATGAGTGTCTGTGAACAAGTTGGAAAAGGACCGAAAACAAG 2397
 Db 241 CAATCTCAACAGATGATGAGTGTCTGTGAACAAGTTGGAAAAGGACCGAAAACAAG 300
 QY 2398 TCGGCAAAAGCCAGAGCATACCCCGAGGAACAGTAACAATGGAAGCACTTCA 2451
 Db 301 TCGGCAAAAGCCAGAGCATACCCCGAGGAACAGTAACAATGGAAGCACTTCA 354

RESULT 36

US-09-922-217-736
 ; Sequence 736, Application US/09922217

; GENERAL INFORMATION:
 ; APPLICANT: Xu, Jianshun
 ; APPLICANT: Lodes, Michael J.
 ; APPLICANT: Secrist, Heather
 ; APPLICANT: Benson, Darin R.
 ; APPLICANT: Meagher, Madeleine Joy
 ; APPLICANT: Stolk, John A.
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Jiang, Yutong
 ; APPLICANT: Smith, Carole Lynn
 ; APPLICANT: King, Gordon E.
 ; APPLICANT: Wang, Aljun
 ; APPLICANT: Clapper, Jonathan D.
 ; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
 ; TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
 ; FILE REFERENCE: 210121.471C13
 ; CURRENT APPLICATION NUMBER: US/09/922,217
 ; CURRENT FILING DATE: 2001-08-03
 ; NUMBER OF SEQ ID NOS: 1124
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 736
 ; LENGTH: 354
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:

;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 14444
;; LENGTH: 496
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC006322.2
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.73
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.55
US-09-864-761-14444

Query Match 10.8%; Score 292; DB 36; Length 496;
Best Local Similarity 100.0%; Pred. No. 7,2e-121; Indels 0; Gaps 0;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2058 AGATCAGATGATGATCATATCATCAGACAGATCAAGCCCTTCTGCTAGCTAGCTTAC 2117
DB 160 AGATCAGATGATGATCATATCATCAGACAGATCAAGCCCTTCTGCTAGCTAGCTTAC 219
QY 2118 AACGAGAGATTCAGGCAATTAACCTTCTGCTAGCTAGCTAGCTTCTTCAATAAG 2177
DB 220 AACGAGAGATTCAGGCAATTAACCTTCTGCTAGCTAGCTAGCTTCTTCAATAAG 279
QY 2178 TTCTTAAGTAACTTGAAGTCAATTCAGACAGAGATTTGGAAGAACTTCTTCAATAAG 2237
DB 280 TTCTTAAGTAACTTGAAGTCAATTCAGACAGAGATTTGGAAGAACTTCTTCAATAAG 339
QY 2238 ATGATGATGAGATGCTCTTAAGACCAAGAAATGTCATAGCATGACACCTAGCCAGA 2297
DB 340 ATGATGATGAGATGCTCTTAAGACCAAGAAATGTCATAGCATGACACCTAGCCAGA 399
QY 2298 AGCTCTGTACAGAGACTTTCATGACGCTCATCAACCCCAATCTCAAC 2349
DB 400 AGCTCTGTACAGAGACTTTCATGACGCTCATCAACCCCAATCTCAAC 451

RESULT 39
US-10-203-135-7281
;; Sequence 7281, Application US/10203135
;; GENERAL INFORMATION:
;; APPLICANT: Molecular Dynamics, Inc.
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: PB 0004 WO 5
;; CURRENT APPLICATION NUMBER: US/10/203,135
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; NUMBER OF SEQ ID NOS: 37012
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 7281
;; LENGTH: 496
;; TYPE: DNA

;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC006322.2
;; FEATURE:
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.73
US-10-203-135-7281

Query Match 10.8%; Score 292; DB 48; Length 496;
Best Local Similarity 100.0%; Pred. No. 7,2e-121; Indels 0; Gaps 0;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2058 AGATCAGATGATGATCATATCATCAGACAGATCAAGCCCTTCTGCTAGCTAGCTTAC 2117
DB 160 AGATCAGATGATGATCATATCATCAGACAGATCAAGCCCTTCTGCTAGCTAGCTTAC 219
QY 2118 AACGAGAGATTCAGGCAATTAACCTTCTGCTAGCTAGCTAGCTTCTTCAATAAG 2177
DB 220 AACGAGAGATTCAGGCAATTAACCTTCTGCTAGCTAGCTAGCTTCTTCAATAAG 279
QY 2178 TTCTTAAGTAACTTGAAGTCAATTCAGACAGAGATTTGGAAGAACTTCTTCAATAAG 2237
DB 280 TTCTTAAGTAACTTGAAGTCAATTCAGACAGAGATTTGGAAGAACTTCTTCAATAAG 339
QY 2238 ATGATGATGAGATGCTCTTAAGACCAAGAAATGTCATAGCATGACACCTAGCCAGA 2297
DB 340 ATGATGATGAGATGCTCTTAAGACCAAGAAATGTCATAGCATGACACCTAGCCAGA 399
QY 2298 AGCTCTGTACAGAGACTTTCATGACGCTCATCAACCCCAATCTCAAC 2349
DB 400 AGCTCTGTACAGAGACTTTCATGACGCTCATCAACCCCAATCTCAAC 451

RESULT 40
US-10-203-136-7203
;; Sequence 7203, Application US/10203136
;; GENERAL INFORMATION:
;; APPLICANT: Molecular Dynamics, Inc.
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: PB 0004 WO 3
;; CURRENT APPLICATION NUMBER: US/10/203,136
;; CURRENT FILING DATE: 2002-08-02
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; NUMBER OF SEQ ID NOS: 38578
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 7203
;; LENGTH: 496
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC006322.2
;; FEATURE:
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.55
US-10-203-136-7203

Query Match 10.8%; Score 292; DB 48; Length 496;

Best Local Similarity 100.0%; Pred. No. 7,2e-121;
Matches 292; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2058 AGATCAGAGTGGATGATCATATCATCAGACAGATCAAGGCTTCTGCTACGTAAGTCTAC 2117
|||||
160 AGATCAGAGTGGATGATCATATCATCAGACAGATCAAGGCTTCTGCTACGTAAGTCTAC 219

2118 AACAGAAAGATTACGGCAATTACCTCTGCCATGCGGCGGAACAATGGGTTCAACAACTC 2177

220 AACAGAAAGATTACGGCAATTACCTCTGCCATGCGGCGGAACAATGGGTTCAACAACTC 279

2178 TTCTTAGGTAACCCCTGGAAGTCATTGACACAGAGCA TTTGGAAGAACTTCTTCATAAAG 2237
|||||
280 TTCTTAGGTAACCCCTGGAAGTCATTGACACAGAGCA TTTGGAAGAACTTCTTCATAAAG 339
|||||

2238 ATGATGATGGAGATGGCTCTTAAGACCAAGAAATGTCCAAATGCGATGACACCTTAGCCAGA 2239
|||||
340 ATGATGATGGAGATGGCTCTTAAGACCAAGAAATGTCCAAATGCGATGACACCTTAGCCAGA 339

2298 AGSTCTGTACAGAGACTTCATGCAGCTCATCAACCACCCCAATCTCAACAC 2349
|||||
400 AGSTCTGTACAGAGACTTCATGCAGCTCATCAACCACCCCAATCTCAACAC 451

RESULT 41
S-09-399-720-14058
Sequence 14058, Application US/09399720
GENERAL INFORMATION

APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
TITLE OF INVENTION: From Various Libraries
U.S. PATENT NO. 2004/0177777

TYPE: DNA
ORGANISM: Homo sapiens
S-09-399-720-14058

Query Match	9.4%	Score 256;	DB 20;	Length 427;
Best Local Similarity	100.0%	Pred. No. 1.5e-104;		
Matches 256; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

2454 AAAATAGAAAGGTAGAAACAGAGGACCCACGAATTTGAGGGGCAACCAGAGTGTCT 2513
102 AAAATAGAAAGGTAGAAACAGAGGACCCACGAATTTGAGAGGGCAACCAGAGTGTCT 161

2514 GAGCTGATTACTCTAGAAACCTCAACAGTAGAACTTGCTAGCAATTAACGTGAA 2513
162 GAGCTGATTACTCTAGAAACCTCAACAGTAGAACTTGCTAGCAATTAACGTGAA 221

2574 AACCAATGCATATACATGACCTTTTTCATGGCATTTATGCGATGTTTACATGGTGG 2633
|||||
222 AAACAAATGCATATACATGACCTTTTTCATGGCATTTATGCGATGTTTACATGGTGG 281

2634 GAAATTGAGCTGAGTTCCACCAATTATTAATTTAAATCCATGATGAATTTCTCTAATAGGC 2693
|||||
282 GAAATTCAGCTGAGTTCCACCAATTATTAATTTAAATCCATGATGAATTTCTCTAATAGGC 341

2694 TTTTTCCTAATACC 2709
 342 TTTTTCCTAATACC 357

RESULT 42
5-09-921-378-14058

SEQUENCE ID: 1000, APPLICATION: 05/055225.78
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained

```

; TITLE OF INVENTION:  From Various Libraries
; FILE REFERENCE:  20411-777

```

```

;      TYPE: DNA
;      ORGANISM: Homo sapiens
US-09-921-378-14058

```

Query Match Score 256: DB 39; Length 427;
Best Local Similarity 100.0%; Pred. No. 1.5e-104;
Matches 256; Conservative 0; Mismatches 0; Gaps 0.

Qy	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500
Db	AAATATGAAAGCTAGAAACAGAGGACCCAGAAATTTGAGAGGACCCAGAGGTGCT	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450																																														

QY	2514	GAGCTGATTACTCTGAAACCTCAACAAGTAGAAACTGCTAGACAATACTGGA	2573
Db	162	GAGCTGATTACTCTGAAACCTCAACAAGTAGAAACTGCTAGACAATACTGGA	221

QY	2574	AAACAAATGCATATACATGAACCTTTTTCATGGCATTATGTCGATGTTTACAAATGGTGG	2633
Nb	222	AAACAAATGCATATACATGAACCTTTTTCATGGCATTATGTCGATGTTTACAAATGGTGG	281

QY 2634 GAAATTCAGCTGATGTCACCAATTAATTAATTCATGAGTAACTTCCATAATAGC 2638

Db 282 GAAATTCAGCTGATGTCACCAATTAATTAATTCATGAGTAACTTCCATAATAGC 341

QY	2694	TTTTTTCCTAATACC	2709
Nb	342	TTTTTTCCTAATACC	357

RESULT 43
JS-09-442-385-770

```

GENERAL INFORMATION:
APPLICANT: Chenchik, Alex
APPLICANT: Lukashov, Matvey
TITLE OF INVENTION: Cancer Array

```

FILE REFERENCE: CLON-U06C1P14
CURRENT APPLICATION NUMBER: US/09/442,385
CURRENT FILING DATE: 1995-11-17
PRIOR APPLICATION NUMBER: 09/053,375

```

; FEATURE: Nucleic Acid Probe
; OTHER INFORMATION:
US-09-442-385-770

```

Query Match	9.4%	Score 254	DB 21	Length 254
Best Local Similarity	100.0%	Pred. No. 1.2e+103		
Matches 254	Conservative 0	Mismatches 0	Indels 0	Gaps 0

QY 1083 TGGCAGGTCCAAATGGCATTTGACACCTCATTTTGGATGAACCTGACGAGTATTTCTTAATGA 1142

Nb 1 TGGCAGGTCCAAATGGCATTTGACACCTCATTTTGGATGAACCTGACGAGTATTTCTTAATGA 60


```
/ CURRENT FILING DATE: 2001-12-20
/ NUMBER OF SEQ ID NOS: 34288
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 18184
/ LENGTH: 164
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO CHR7.1
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.59
/ OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
/ OTHER INFORMATION: SWISSPROT HIT: BF667677.1, EVALUE 4.00e-29
/ OTHER INFORMATION: EST HUMAN HIT: BF667677.1, EVALUE 9.00e-81
/ OTHER INFORMATION: NT HIT: g116162409, EVALUE 1.00e-85
US-10-029-386-18184

Query Match          5.9%; Score 161; DB 44; Length 164;
Best Local Similarity 100.0%; Pred. No. 1.7e-61;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1691 CAGCAACAATAATATATGTTCAAGCGCTGGGCTTCCCTTACACCGGTGT 1750
    |||
DB 164 CAGCAACAATAATATATGTTCAAGCGCTGGGCTTCCCTTACACCGGTGT 105

QY 1751 GATATTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGACCTTACTGTGCTGG 1810
    |||
DB 104 GATATTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGACCTTACTGTGCTGG 45

QY 1811 GATGTTCTGCAGTGTCTGCGTATTTTCCCATGCAAGAAG 1851
    |||
DB 44 GATGTTCTGCAGTGTCTGCGTATTTTCCCATGCAAGAAG 4

RESULT 48
US-10-029-386-4484/c
/ Sequence 4484, Application US/10029386
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
/ FILE REFERENCE: AEOMICA-X-2
/ CURRENT APPLICATION NUMBER: US/10/029,386
/ CURRENT FILING DATE: 2001-12-20
/ NUMBER OF SEQ ID NOS: 34288
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 4484
/ LENGTH: 591
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO CHR7.1
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.59
/ OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.4
/ OTHER INFORMATION: NT HIT: g116162409, EVALUE 4.00e-85
/ OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 2.00e-28
/ OTHER INFORMATION: EST_HUMAN HIT: BF667677.1, EVALUE 4.00e-80
US-10-029-386-4484

Query Match          5.9%; Score 161; DB 44; Length 591;
Best Local Similarity 100.0%; Pred. No. 1.6e-61;
Matches 161; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1691 CAGCAACAATAATATATGTTCAAGCGCTGGGCTTCCCTTACACCGGTGT 1750
    |||
DB 311 CAGCAACAATAATATATGTTCAAGCGCTGGGCTTCCCTTACACCGGTGT 252

QY 1751 GATATTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGACCTTACTGTGCTGG 1810
```

```
DB 251 GATATTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGACCTTACTGTGCTGG 192
    |||
QY 1811 GATGTTCTGCAGTGTCTGCGTATTTTCCCATGCAAGAAG 1851
    |||
DB 191 GATGTTCTGCAGTGTCTGCGTATTTTCCCATGCAAGAAG 151

RESULT 49
PCT-US01-00663-20125/c
/ Sequence 20125, Application PC/TUS0100663
/ GENERAL INFORMATION:
/ APPLICANT: Molecular Dynamics, Inc.
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ FILE REFERENCE: PB 0004 WO 7
/ CURRENT APPLICATION NUMBER: PCT/US01/00663
/ CURRENT FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 04 February 2000 (04.02.00)
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 26 May 2000 (26.05.00)
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 03 August 2000 (03.08.00)
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 03 October 2000 (03.10.00)
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 27 September 2000 (27.09.00)
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 21 September 2000 (21.09.00)
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 30 June 2000 (30.06.00)
/ NUMBER OF SEQ ID NOS: 38837
/ SOFTWARE: Molecular Dynamics Sequence Listing Engine
/ SEQ ID NO 20125
/ LENGTH: 172
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC004848.1
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
/ OTHER INFORMATION: NT HIT: g11421514, EVALUE 8.00e-89
/ OTHER INFORMATION: EST_HUMAN HIT: BE694785.1, EVALUE 2.30e+00
/ OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
PCT-US01-00663-20125

Query Match          5.9%; Score 160; DB 1; Length 172;
Best Local Similarity 100.0%; Pred. No. 4.8e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAAATGTTGGAATCCAAATATGATCATCTTCAATGCGCTTGGCCAAAGCTCAGTTA 369
    |||
DB 160 ABAATGTTGGAATCCAAATATGATCATCTTCAATGCGCTTGGCCAAAGCTCAGTTA 101

QY 370 TATATCTTCTCTTTTGATGAGAAAGATGAGCTGTATGTGAGCAAGAGATCAAT 429
    |||
DB 100 TATATCTTCTCTTTTGATGAGAAAGATGAGCTGTATGTGAGCAAGAGATCAAT 41

QY 430 ATTTCATTGCAGCTGTTAATATATCAAGATTTTCAAG 469
    |||
DB 40 ATTTCATTGCAGCTGTTAATATATCAAGATTTTCAAG 1

RESULT 50
PCT-US01-00663-20125/c
/ Sequence 20125, Application PC/TUS0100663
/ GENERAL INFORMATION:
/ APPLICANT: Molecular Dynamics, Inc.
/ APPLICANT: Penn, Sharon G.
```

```
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 7
CURRENT APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38837
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 20125
LENGTH: 172
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
OTHER INFORMATION: NT HIT: G11421514, EVALUE 8.00e-89
OTHER INFORMATION: EST HUMAN HIT: BE694785.1, EVALUE 2.30e+00
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
PCT-US01-00663-20125

Query Match          5.9%; Score 160; DB 2; Length 172;
Best Local Similarity 100.0%; Pred. No. 4.8e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGAATCCAAATGATGATCACTTTCAATGGCTTGGCCAAAGCTCCAGTTA 369
DB 160 AGAATGTTGAATCCAAATGATGATCACTTTCAATGGCTTGGCCAAAGCTCCAGTTA 101

QY 370 TCATACCTTCTTTTGGATGAGGAACGAGTAGGTGATGTTGGAGCAAGATCAAT 429
DB 100 TCATACCTTCTTTTGGATGAGGAACGAGTAGGTGATGTTGGAGCAAGATCAAT 41

QY 430 ATTTTCATTCGACCTGGTTAATATCAAGATTTTCAAAAG 469
DB 40 ATTTTCATTCGACCTGGTTAATATCAAGATTTTCAAAAG 1

RESULT 51
US-09-864-761-27943/c
Sequence 27943, Application US/09864761
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aecmics-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
```

```
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 27943
LENGTH: 172
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
OTHER INFORMATION: NT HIT: G11421514, EVALUE 8.00e-89
OTHER INFORMATION: EST HUMAN HIT: BE694785.1, EVALUE 2.30e+00
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
US-09-864-761-27943

Query Match          5.9%; Score 160; DB 36; Length 172;
Best Local Similarity 100.0%; Pred. No. 4.8e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGAATCCAAATGATGATCACTTTCAATGGCTTGGCCAAAGCTCCAGTTA 369
DB 160 AGAATGTTGAATCCAAATGATGATCACTTTCAATGGCTTGGCCAAAGCTCCAGTTA 101

QY 370 TCATACCTTCTTTTGGATGAGGAACGAGTAGGTGATGTTGGAGCAAGATCAAT 429
DB 100 TCATACCTTCTTTTGGATGAGGAACGAGTAGGTGATGTTGGAGCAAGATCAAT 41

QY 430 ATTTTCATTCGACCTGGTTAATATCAAGATTTTCAAAAG 469
DB 40 ATTTTCATTCGACCTGGTTAATATCAAGATTTTCAAAAG 1

RESULT 52
US-10-182-993-19495/c
Sequence 19495, Application US/10182993
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
```

```
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 2 10/182,993
CURRENT FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 19495
LENGTH: 172
TYPE: DNA
ORGANISM: Homo sapiens
OTHER INFORMATION: MAP TO AC004848.1
FEATURE:
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
FEATURE:
OTHER INFORMATION: NT HIT: g111421514, EVALUE 8.00e-89
FEATURE:
OTHER INFORMATION: EST_HUMAN HIT: BE694785.1, EVALUE 2.30e+00
FEATURE:
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
US-10-182-993-19495

Query Match
Best Local Similarity 100.0%; Pred. No. 4.8e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCACATGTGATCACTTTCATGGCTTGCCACAGCTCCAGTTA 369
DB 160 AGAATGTTGGAATCCACATGTGATCACTTTCATGGCTTGCCACAGCTCCAGTTA 101
QY 370 TCATACCTTCCTTTGGATGAGAACGAGTAGGCTGTATGTTGGACCAAGATCAAT 429
DB 100 TCATACCTTCCTTTGGATGAGAACGAGTAGGCTGTATGTTGGACCAAGATCAAT 41
QY 430 ATTTTCATTGACCTGTTAATATCAAGGATTTTCAAAAG 469
DB 40 ATTTTCATTGACCTGTTAATATCAAGGATTTTCAAAAG 1

RESULT 53
US-10-203-134-20052/c
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 6
CURRENT FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
```

```
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 20052
LENGTH: 172
TYPE: DNA
ORGANISM: Homo sapiens
OTHER INFORMATION: MAP TO AC004848.1
FEATURE:
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
FEATURE:
OTHER INFORMATION: NT HIT: g111421514, EVALUE 8.00e-89
FEATURE:
OTHER INFORMATION: EST_HUMAN HIT: BE694785.1, EVALUE 2.30e+00
FEATURE:
OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
US-10-203-134-20052

Query Match
Best Local Similarity 100.0%; Pred. No. 4.8e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCACATGTGATCACTTTCATGGCTTGCCACAGCTCCAGTTA 369
DB 160 AGAATGTTGGAATCCACATGTGATCACTTTCATGGCTTGCCACAGCTCCAGTTA 101
QY 370 TCATACCTTCCTTTGGATGAGAACGAGTAGGCTGTATGTTGGACCAAGATCAAT 429
DB 100 TCATACCTTCCTTTGGATGAGAACGAGTAGGCTGTATGTTGGACCAAGATCAAT 41
QY 430 ATTTTCATTGACCTGTTAATATCAAGGATTTTCAAAAG 469
DB 40 ATTTTCATTGACCTGTTAATATCAAGGATTTTCAAAAG 1

RESULT 54
US-10-203-135-19754/c
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 5
CURRENT FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
```

```
/ NUMBER OF SEQ ID NOS: 37012
/ SOFTWARE: Molecular Dynamics Sequence Listing Engine
/ SEQ ID NO 19754
/ LENGTH: 172
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC004848.1
/ FEATURE:
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
/ FEATURE:
/ OTHER INFORMATION: NT HIT: g111421514, EVALUE 8.00e-89
/ FEATURE:
/ OTHER INFORMATION: EST_HUMAN HIT: BE694785.1, EVALUE 2.30e+00
/ FEATURE:
/ OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
US-10-203-135-19754

Query Match
Best Local Similarity 100.0%; Pred. No. 4.8e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 310 AGAATGTTGGAATCCAAATGATGATCATCTTCAATGGCTGGCCAAAGCTCCAGTTA 369
Db 160 AGAATGTTGGAATCCAAATGATGATCATCTTCAATGGCTGGCCAAAGCTCCAGTTA 101

Qy 370 TCATACCTTCTTTTGGATGAGAACGAGTAGGCTGTATGTTGGAGCAAGATCACAT 429
Db 100 TCATACCTTCTTTTGGATGAGAACGAGTAGGCTGTATGTTGGAGCAAGATCACAT 41

Qy 430 ATTTTCATTGACCTGGTTATATATCAAGGATTTTCAAAAG 469
Db 40 ATTTTCATTGACCTGGTTATATATCAAGGATTTTCAAAAG 1

RESULT 55
US-10-203-136-20171/c
/ Sequence 20171, Application US/10203136
/ GENERAL INFORMATION:
/ APPLICANT: Molecular Dynamics, Inc.
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ FILE REFERENCE: PB 0004 WO 3
/ CURRENT APPLICATION NUMBER: US/10/203,136
/ PRIOR FILING DATE: 2002-08-02
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 04 February 2000 (04.02.00)
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 26 May 2000 (26.05.00)
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 03 August 2000 (03.08.00)
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 03 October 2000 (03.10.00)
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 27 September 2000 (27.09.00)
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 21 September 2000 (21.09.00)
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 30 June 2000 (30.06.00)
/ NUMBER OF SEQ ID NOS: 38578
/ SOFTWARE: Molecular Dynamics Sequence Listing Engine
/ SEQ ID NO 20171
/ LENGTH: 172
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC004848.1
/ FEATURE:
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
```

```
/ FEATURE:
/ OTHER INFORMATION: NT HIT: g111421514, EVALUE 8.00e-89
/ FEATURE:
/ OTHER INFORMATION: EST_HUMAN HIT: BE694785.1, EVALUE 2.30e+00
/ FEATURE:
/ OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
US-10-203-136-20171

Query Match
Best Local Similarity 100.0%; Pred. No. 4.8e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 310 AGAATGTTGGAATCCAAATGATGATCATCTTCAATGGCTGGCCAAAGCTCCAGTTA 369
Db 160 AGAATGTTGGAATCCAAATGATGATCATCTTCAATGGCTGGCCAAAGCTCCAGTTA 101

Qy 370 TCATACCTTCTTTTGGATGAGAACGAGTAGGCTGTATGTTGGAGCAAGATCACAT 429
Db 100 TCATACCTTCTTTTGGATGAGAACGAGTAGGCTGTATGTTGGAGCAAGATCACAT 41

Qy 430 ATTTTCATTGACCTGGTTATATATCAAGGATTTTCAAAAG 469
Db 40 ATTTTCATTGACCTGGTTATATATCAAGGATTTTCAAAAG 1

RESULT 56
US-10-203-137-20125/c
/ Sequence 20125, Application US/10203137
/ GENERAL INFORMATION:
/ APPLICANT: Molecular Dynamics, Inc.
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ FILE REFERENCE: PB 0004 WO 7
/ CURRENT APPLICATION NUMBER: US/10/203,137
/ PRIOR FILING DATE: 2002-08-02
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 04 February 2000 (04.02.00)
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 26 May 2000 (26.05.00)
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 03 August 2000 (03.08.00)
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 03 October 2000 (03.10.00)
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 27 September 2000 (27.09.00)
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 21 September 2000 (21.09.00)
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 30 June 2000 (30.06.00)
/ NUMBER OF SEQ ID NOS: 38837
/ SOFTWARE: Molecular Dynamics Sequence Listing Engine
/ SEQ ID NO 20125
/ LENGTH: 172
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC004848.1
/ FEATURE:
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
/ FEATURE:
/ OTHER INFORMATION: NT HIT: g111421514, EVALUE 8.00e-89
/ FEATURE:
/ OTHER INFORMATION: EST_HUMAN HIT: BE694785.1, EVALUE 2.30e+00
/ FEATURE:
/ OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 5.00e-25
US-10-203-137-20125

Query Match
Best Local Similarity 100.0%; Pred. No. 4.8e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
5.9%; Score 160; DB 48; Length 172;
100.0%; Pred. No. 4.8e-61;
5.9%; Score 160; DB 48; Length 172;
100.0%; Pred. No. 4.8e-61;
```


;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; NUMBER OF SEQ ID NOS: 38837
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 7072
;; LENGTH: 484
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC004848.1
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
PCT-US01-00663-7072

Query Match 5.9%; Score 160; DB 1; Length 484;
Best Local Similarity 100.0%; Pred. No. 4,6e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 310 AGAATGTTGGAATCAACAATGTGATCACTTCAATGGCTGGCCAAAGCTCCAGTTA 369
Db 458 AGAATGTTGGAATCAACAATGTGATCACTTCAATGGCTGGCCAAAGCTCCAGTTA 399
Qy 370 TCATACCTTCTTTGGATGAGAAAGAGTGTATGTTGGAGCAAGATCACAT 429
Db 398 TCATACCTTCTTTGGATGAGAAAGAGTGTATGTTGGAGCAAGATCACAT 339
Qy 430 ATTTCATTGCACTGGTTAATATCAAGATTTTCAAAG 469
Db 338 ATTTCATTGCACTGGTTAATATCAAGATTTTCAAAG 299

RESULT 60

PCT-US01-00663-7072/c
;; Sequence 7072, Application PC/TUS0100663
;; GENERAL INFORMATION:
;; APPLICANT: Molecular Dynamics, Inc.
;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: PB 0004 WO 7
;; CURRENT APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 04 February 2000 (04.02.00)
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 26 May 2000 (26.05.00)
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 03 August 2000 (03.08.00)
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 03 October 2000 (03.10.00)
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 27 September 2000 (27.09.00)
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 21 September 2000 (21.09.00)
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 30 June 2000 (30.06.00)
;; NUMBER OF SEQ ID NOS: 38837
;; SOFTWARE: Molecular Dynamics Sequence Listing Engine
;; SEQ ID NO 7072
;; LENGTH: 484
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC004848.1
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
PCT-US01-00663-7072

Query Match 5.9%; Score 160; DB 2; Length 484;
Best Local Similarity 100.0%; Pred. No. 4,6e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 310 AGAATGTTGGAATCAACAATGTGATCACTTCAATGGCTGGCCAAAGCTCCAGTTA 369
Db 458 AGAATGTTGGAATCAACAATGTGATCACTTCAATGGCTGGCCAAAGCTCCAGTTA 399
Qy 370 TCATACCTTCTTTGGATGAGAAAGAGTGTATGTTGGAGCAAGATCACAT 429
Db 398 TCATACCTTCTTTGGATGAGAAAGAGTGTATGTTGGAGCAAGATCACAT 339
Qy 430 ATTTCATTGCACTGGTTAATATCAAGATTTTCAAAG 469
Db 338 ATTTCATTGCACTGGTTAATATCAAGATTTTCAAAG 299

RESULT 61

US-09-864-761-11363/c
;; Sequence 11363, Application US/09864761
;; GENERAL INFORMATION:

;; APPLICANT: Penn, Sharon G.
;; APPLICANT: Rank, David R.
;; APPLICANT: Hanzel, David K.
;; APPLICANT: Chen, Wensheng
;; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;; FILE REFERENCE: Aecomica-X-1
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; PRIOR FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263.6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 11363
;; LENGTH: 484
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:

OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
US-09-761-11363

Query Match 5.9%; Score 160; DB 36; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCAAATGATGATCACTTCAATGGCTGGCCAAAGCTCCAGTTA 369
DB 458 AGAATGTTGGAATCCAAATGATGATCACTTCAATGGCTGGCCAAAGCTCCAGTTA 399
QY 370 TCATACCTCTCTTTGGATGAGAAACGAGTGGTATGTTGAGCAAAAGATCAAT 429
DB 398 TCATACCTCTCTTTGGATGAGAAACGAGTGGTATGTTGAGCAAAAGATCAAT 339
QY 430 ATTTCAATCGACCTGGTTAATATCAAGATTTTCAAAAG 469
DB 338 ATTTCAATCGACCTGGTTAATATCAAGATTTTCAAAAG 299

RESULT 62
US-10-182-993-6833/c

Sequence 6833, Application US/10182993
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 2
CURRENT FILING DATE: 2002-08-02
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 37811
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 6833
LENGTH: 484
TYPE: DNA
ORGANISM: Homo sapiens
OTHER INFORMATION: MAP TO AC004848.1
FEATURE:
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
US-10-182-993-6833

Query Match 5.9%; Score 160; DB 47; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCAAATGATGATCACTTCAATGGCTGGCCAAAGCTCCAGTTA 369
DB 458 AGAATGTTGGAATCCAAATGATGATCACTTCAATGGCTGGCCAAAGCTCCAGTTA 399

QY 370 TCATACCTCTCTTTGGATGAGAAACGAGTGGTATGTTGAGCAAAAGATCAAT 429
DB 398 TCATACCTCTCTTTGGATGAGAAACGAGTGGTATGTTGAGCAAAAGATCAAT 339
QY 430 ATTTCAATCGACCTGGTTAATATCAAGATTTTCAAAAG 469
DB 338 ATTTCAATCGACCTGGTTAATATCAAGATTTTCAAAAG 299

RESULT 63
US-10-203-134-7118/c

Sequence 7118, Application US/10203134
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 6
CURRENT FILING DATE: 2002-08-02
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38628
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 7118
LENGTH: 484
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
US-10-203-134-7118

Query Match 5.9%; Score 160; DB 48; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCAAATGATGATCACTTCAATGGCTGGCCAAAGCTCCAGTTA 369
DB 458 AGAATGTTGGAATCCAAATGATGATCACTTCAATGGCTGGCCAAAGCTCCAGTTA 399
QY 370 TCATACCTCTCTTTGGATGAGAAACGAGTGGTATGTTGAGCAAAAGATCAAT 429
DB 398 TCATACCTCTCTTTGGATGAGAAACGAGTGGTATGTTGAGCAAAAGATCAAT 339
QY 430 ATTTCAATCGACCTGGTTAATATCAAGATTTTCAAAAG 469
DB 338 ATTTCAATCGACCTGGTTAATATCAAGATTTTCAAAAG 299

RESULT 64
US-10-203-135-7342/c

Sequence 7342, Application US/10203135
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.


```
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME- DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 5
CURRENT APPLICATION NUMBER: US/10/203,135
PRIOR FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 37012
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 7342
LENGTH: 484
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
FEATURE:
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
US-10-203-135-7342
```

```
Query Match          5.9%; Score 160; DB 48; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 310 AGAATGTTGGATCAACAAATGATGATCCTTCAATGGCTGGCCAAAGCTCCAGTTA 369
DB 458 AGAATGTTGGATCAACAAATGATGATCCTTCAATGGCTGGCCAAAGCTCCAGTTA 399
QY 370 TCATACCTTCTTTGGATGAGAAAGAGTAGAGCTGTATGTTGGAGCAAGATCACAT 429
DB 398 TCATACCTTCTTTGGATGAGAAAGAGTAGAGCTGTATGTTGGAGCAAGATCACAT 339
QY 430 ATTTTCATTCGACCTGTTAATATCAAGGATTTTCAAAAG 469
DB 338 ATTTTCATTCGACCTGTTAATATCAAGGATTTTCAAAAG 299
```

```
RESULT 65
US-10-203-136-7262/c
Sequence 7262, Application US/10203136
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME- DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 3
CURRENT APPLICATION NUMBER: US/10/203,136
PRIOR FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
```

```
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38578
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 7262
LENGTH: 484
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
FEATURE:
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
US-10-203-136-7262
```

```
Query Match          5.9%; Score 160; DB 48; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 310 AGAATGTTGGATCAACAAATGATGATCCTTCAATGGCTGGCCAAAGCTCCAGTTA 369
DB 458 AGAATGTTGGATCAACAAATGATGATCCTTCAATGGCTGGCCAAAGCTCCAGTTA 399
QY 370 TCATACCTTCTTTGGATGAGAAAGAGTAGAGCTGTATGTTGGAGCAAGATCACAT 429
DB 398 TCATACCTTCTTTGGATGAGAAAGAGTAGAGCTGTATGTTGGAGCAAGATCACAT 339
QY 430 ATTTTCATTCGACCTGTTAATATCAAGGATTTTCAAAAG 469
DB 338 ATTTTCATTCGACCTGTTAATATCAAGGATTTTCAAAAG 299
```

```
RESULT 66
US-10-203-137-7072/c
Sequence 7072, Application US/10203137
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharron G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME- DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 7
CURRENT APPLICATION NUMBER: US/10/203,137
PRIOR FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38837
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 7072
LENGTH: 484
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004848.1
FEATURE:
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
```

US-10-203-137-7072

Query Match 5.9%; Score 160; DB 48; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCACATGTGATCACTTCAATGGCTTGCCCAACGCTCCAGTTA 369
DB 458 AGAATGTTGGAATCCACATGTGATCACTTCAATGGCTTGCCCAACGCTCCAGTTA 399
QY 370 TCATACCTTCCTTTGGATGAGAACGAGTAGCTGTATGTTGAGCAAGATCAAT 429
DB 398 TCATACCTTCCTTTGGATGAGAACGAGTAGCTGTATGTTGAGCAAGATCAAT 359
QY 430 ATTTTCATTGACCTGGTTAATATCAAGATTTTCAAAAG 469
DB 338 ATTTTCATTGACCTGGTTAATATCAAGATTTTCAAAAG 299

RESULT 67

US-10-203-139-7012/c

Sequence 7012, Application US/10203139

GENERAL INFORMATION:

APPLICANT: Molecular Dynamics, Inc.

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN FETAL LIVER

FILE REFERENCE: PB 0004 WO 4

CURRENT APPLICATION NUMBER: US/10/203,139

CURRENT FILING DATE: 2002-08-02

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 04 February 2000 (04.02.00)

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 26 May 2000 (26.05.00)

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 03 August 2000 (03.08.00)

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 03 October 2000 (03.10.00)

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 27 September 2000 (27.09.00)

PRIOR APPLICATION NUMBER: US 60/234,687

PRIOR FILING DATE: 21 September 2000 (21.09.00)

PRIOR APPLICATION NUMBER: US 09/608,408

PRIOR FILING DATE: 30 June 2000 (30.06.00)

NUMBER OF SEQ ID NOS: 37156

SOFTWARE: Molecular Dynamics Sequence Listing Engine

SEQ ID NO 7012

LENGTH: 484

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO AC004848.1

FEATURE:

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1

US-10-203-139-7012

Query Match 5.9%; Score 160; DB 48; Length 484;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCACATGTGATCACTTCAATGGCTTGCCCAACGCTCCAGTTA 369
DB 458 AGAATGTTGGAATCCACATGTGATCACTTCAATGGCTTGCCCAACGCTCCAGTTA 399
QY 370 TCATACCTTCCTTTGGATGAGAACGAGTAGCTGTATGTTGAGCAAGATCAAT 429
DB 398 TCATACCTTCCTTTGGATGAGAACGAGTAGCTGTATGTTGAGCAAGATCAAT 339
QY 430 ATTTTCATTGACCTGGTTAATATCAAGATTTTCAAAAG 469
DB 338 ATTTTCATTGACCTGGTTAATATCAAGATTTTCAAAAG 469

DB 338 ATTTTCATTGACCTGGTTAATATCAAGATTTTCAAAAG 299

RESULT 68

US-10-029-386-2791

Sequence 2791, Application US/10029386

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G

TITLE OF INVENTION: EXPRESSION ANALYSIS TWO

FILE REFERENCE: AEOMICA-X-2

CURRENT APPLICATION NUMBER: US/10/029,386

CURRENT FILING DATE: 2001-12-20

NUMBER OF SEQ ID NOS: 34288

SOFTWARE: Annomax Sequence Listing Engine vers. 1.1

SEQ ID NO 2791

LENGTH: 526

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: MAP TO CHR7.1

OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.5

OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.68

OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.64

OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.7

OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.69

OTHER INFORMATION: NT HIT: g15174672, EVALUE 1.00e-84

OTHER INFORMATION: EST HUMAN HIT: B1552379.1, EVALUE 2.90e+00

OTHER INFORMATION: SWISSPROT HIT: Q14563, EVALUE 3.00e-25

US-10-029-386-2791

Query Match 5.9%; Score 160; DB 44; Length 526;
Best Local Similarity 100.0%; Pred. No. 4.6e-61;
Matches 160; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 310 AGAATGTTGGAATCCACATGTGATCACTTCAATGGCTTGCCCAACGCTCCAGTTA 369
DB 161 AGAATGTTGGAATCCACATGTGATCACTTCAATGGCTTGCCCAACGCTCCAGTTA 220
QY 370 TCATACCTTCCTTTGGATGAGAACGAGTAGCTGTATGTTGAGCAAGATCAAT 429
DB 221 TCATACCTTCCTTTGGATGAGAACGAGTAGCTGTATGTTGAGCAAGATCAAT 280
QY 430 ATTTTCATTGACCTGGTTAATATCAAGATTTTCAAAAG 469
DB 281 ATTTTCATTGACCTGGTTAATATCAAGATTTTCAAAAG 320

RESULT 69

US-09-864-761-11810

Sequence 11810, Application US/09864761

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR

TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

FILE REFERENCE: Aeomica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

```

; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO: 11810
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006322.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.46
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.54
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56
; US-09-864-761-11810

Query Match
Best Local Similarity 5.3%; Score 144; DB 36; Length 446;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1916 GATAATCACCATTGGCCAGAGCCCTGAGAGAGATCATCTATGTGTGAGAAATGTAGC 1975
Db 277 GATAATCACCATTGGCCAGAGCCCTGAGAGAGATCATCTATGTGTGAGAAATGTAGC 336
Qy 1976 ACATTTTGAATGAGTCGAGTCGGAAGTCGAGAGAGCGCTGTCTATTGGCAATTCAGAG 2035
Db 337 ACATTTTGAATGAGTCGAGTCGGAAGTCGAGAGAGCGCTGTCTATTGGCAATTCAGAG 396
Qy 2036 CGAATGAAGAGCGGAAAGAGAG 2059
Db 397 CGAATGAAGAGCGGAAAGAGAG 420

RESULT 70
US-10-182-993-7713
; Sequence 7713, Application US/10182993
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME- DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 2
; FILE REFERENCE: ANALYSIS OF GENE EXPRESSION IN HUMAN BRAIN
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US/10/182,993
; PRIOR FILING DATE: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
```

```

; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 37811
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO: 7713
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006322.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.56
; US-10-182-993-7713

Query Match
Best Local Similarity 5.3%; Score 144; DB 47; Length 446;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1916 GATAATCACCATTGGCCAGAGCCCTGAGAGAGATCATCTATGTGTGAGAAATGTAGC 1975
Db 277 GATAATCACCATTGGCCAGAGCCCTGAGAGAGATCATCTATGTGTGAGAAATGTAGC 336
Qy 1976 ACATTTTGAATGAGTCGAGTCGGAAGTCGAGAGAGCGCTGTCTATTGGCAATTCAGAG 2035
Db 337 ACATTTTGAATGAGTCGAGTCGGAAGTCGAGAGAGCGCTGTCTATTGGCAATTCAGAG 396
Qy 2036 CGAATGAAGAGCGGAAAGAGAG 2059
Db 397 CGAATGAAGAGCGGAAAGAGAG 420

RESULT 71
US-10-203-135-8417
; Sequence 8417, Application US/10203135
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME- DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: PB 0004 WO 5
; FILE REFERENCE: ANALYSIS OF GENE EXPRESSION IN HUMAN LUNG
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: US/10/203,135
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 37012
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO: 8417
; LENGTH: 446
; TYPE: DNA
```

ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006322.2
FEATURE:
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.54
US-10-203-135-8417

Query Match
Best Local Similarity 100.0%; Pred. No. 8.3e-54;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1916 GATATCACCATGCGCAGCCCTGAAGAGAGATCATCTATGTGTAGAGATTAGTAC 1975
DB 277 GATATCACCATGCGCAGCCCTGAAGAGAGATCATCTATGTGTAGAGATTAGTAC 336
QY 1976 ACATTTTGAATGCGAGTGCAGAGAGCGCTGCTATTGGCAATTCCAGAG 2035
DB 337 ACATTTTGAATGCGAGTGCAGAGAGCGCTGCTATTGGCAATTCCAGAG 396
QY 2036 CGAATGAAGCGCAAGAGAGAG 2059
DB 397 CGAATGAAGCGCAAGAGAGAG 420

RESULT 72
US-10-203-136-8331
Sequence 8331, Application US/10203136
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: PB 0004 WO 3
CURRENT FILING DATE: 2002-08-02
PRIOR FILING DATE: 2002-08-02
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/226,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38578
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 8331
LENGTH: 446
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC006322.2
FEATURE:
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.46
US-10-203-136-8331

Query Match
Best Local Similarity 100.0%; Pred. No. 8.3e-54;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1916 GATATCACCATGCGCAGCCCTGAAGAGAGATCATCTATGTGTAGAGATTAGTAC 1975
DB 277 GATATCACCATGCGCAGCCCTGAAGAGAGATCATCTATGTGTAGAGATTAGTAC 336
QY 1976 ACATTTTGAATGCGAGTGCAGAGAGCGCTGCTATTGGCAATTCCAGAG 2035

DB 337 ACATTTTGAATGCGAGTGCAGAGAGCGCTGCTATTGGCAATTCCAGAG 396
QY 2036 CGAATGAAGCGCAAGAGAGAG 2059
DB 397 CGAATGAAGCGCAAGAGAGAG 420

RESULT 73
US-60-160-203-771/c
Sequence 771, Application US/60160203
GENERAL INFORMATION:
APPLICANT: BONAZZI, VIVIAN
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
FILE REFERENCE: C1000116
CURRENT APPLICATION NUMBER: US/60/160,203
CURRENT FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 6374
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 771
LENGTH: 574
TYPE: DNA
ORGANISM: HUMAN
US-60-160-203-771

Query Match
Best Local Similarity 100.0%; Pred. No. 8.2e-54;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 866 GATCCAAAGTTTATTAGTCCCACTCTCATCTCAGAGAGTACATCTGTAAGATGACAA 925
DB 455 GATCCAAAGTTTATTAGTCCCACTCTCATCTCAGAGAGTACATCTGTAAGATGACAA 396
QY 926 GATATCTTTTCTTCGCGAAGATGATGAGAACTCTGGAAGAACTACTCAG 985
DB 395 GATATCTTTTCTTCGCGAAGATGATGAGAACTCTGGAAGAACTACTCAG 336
QY 986 GCTAGATAGTGCATATGCAAG 1009
DB 335 GCTAGATAGTGCATATGCAAG 312

RESULT 74
US-60-160-203-838
Sequence 838, Application US/60160203
GENERAL INFORMATION:
APPLICANT: BONAZZI, VIVIAN
TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
FILE REFERENCE: C1000116
CURRENT APPLICATION NUMBER: US/60/160,203
CURRENT FILING DATE: 1999-10-19
NUMBER OF SEQ ID NOS: 6374
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 838
LENGTH: 728
TYPE: DNA
ORGANISM: HUMAN
US-60-160-203-838

Query Match
Best Local Similarity 100.0%; Pred. No. 8.1e-54;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 866 GATCCAAAGTTTATTAGTCCCACTCTCATCTCAGAGAGTACATCTGTAAGATGACAA 925
DB 322 GATCCAAAGTTTATTAGTCCCACTCTCATCTCAGAGAGTACATCTGTAAGATGACAA 381
QY 926 GATATCTTTTCTTCGCGAAGATGATGAGAACTCTGGAAGAACTACTCAG 985

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 19:48:41 ; Search time 327 Seconds
(without alignments)
4132.889 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709
Sequence: 1 aatcttattatcgcagtc.....agcctttttccctaacc 2709

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 424677 seqs, 249437934 residues

Word size : 0

Total number of hits satisfying chosen parameters: 849354

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 100 summaries

Database :

Pending_Patents_NA_New.*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	26	1.0	1313	US-10-293-244-975	Sequence 975, App
2	23	0.8	7476	PCT-US03-09662-4	Sequence 4, Appl
3	23	0.8	7494	PCT-US03-09662-3	Sequence 3, Appl
4	23	0.8	7617	PCT-US03-09662-1	Sequence 1, Appl
5	23	0.8	7626	PCT-US03-09662-2	Sequence 2, Appl
6	21	0.8	906	US-10-293-244-459	Sequence 459, App
7	21	0.8	1026	US-10-357-930-22912	Sequence 22912, A
8	21	0.8	1026	US-10-357-930-28772	Sequence 28772, A
9	21	0.8	1181	US-10-105-837-765	Sequence 765, App
10	21	0.8	1538	US-10-293-244-765	Sequence 664, App
11	21	0.8	1572	US-10-105-837-664	Sequence 27697, A
12	21	0.8	2028	US-10-357-930-27697	Sequence 54, Appl
13	21	0.8	3091	US-09-663-481-2	Sequence 54, Appl
14	21	0.8	5511	US-10-229-541A-54	Sequence 55, Appl
15	21	0.8	5517	US-10-229-541A-79	Sequence 79, Appl
16	21	0.8	5635	US-10-229-541A-55	Sequence 56, Appl
17	21	0.8	6299	US-10-229-541A-56	Sequence 56, Appl
18	21	0.8	6299	US-10-229-541A-56	Sequence 56, Appl
19	21	0.8	6423	US-10-229-541A-56	Sequence 57, Appl
20	21	0.8	6423	US-10-229-541A-57	Sequence 57, Appl
21	20	0.7	598	US-10-357-930-50683	Sequence 50683, A
22	20	0.7	1168	US-10-105-837-562	Sequence 562, App
23	20	0.7	1423	US-10-391-363A-59	Sequence 30, Appl
24	20	0.7	1498	US-10-115-831-30	Sequence 30, Appl
25	20	0.7	1522	US-10-115-831-102	Sequence 102, App
26	20	0.7	1906	US-10-293-244-480	Sequence 480, App

27	20	0.7	1920	US-10-105-837-7	Sequence 7, Appl
28	20	0.7	2225	US-10-105-837-1046	Sequence 1046, App
29	20	0.7	2946	US-10-357-930-22213	Sequence 22213, A
30	20	0.7	2946	US-10-357-930-23051	Sequence 23051, A
31	20	0.7	2946	US-10-357-930-28070	Sequence 28070, A
32	20	0.7	2946	US-10-357-930-28916	Sequence 28916, A
33	20	0.7	3336	US-09-654-936A-134	Sequence 134, App
34	20	0.7	3336	US-10-293-244-900	Sequence 900, App
35	20	0.7	13831263	US-09-947-914-41	Sequence 41, Appl
36	19	0.7	433	US-10-273-573-1607	Sequence 1607, App
37	19	0.7	470	US-10-085-783A-26572	Sequence 26572, A
38	19	0.7	479	US-10-085-783A-39690	Sequence 39690, A
39	19	0.7	530	US-10-357-930-21683	Sequence 21683, A
40	19	0.7	530	US-10-357-930-27527	Sequence 27527, A
41	19	0.7	571	US-10-105-837-920	Sequence 920, App
42	19	0.7	786	US-10-273-573-4669	Sequence 4669, App
43	19	0.7	1020	US-10-357-930-22333	Sequence 22333, A
44	19	0.7	1020	US-10-357-930-24125	Sequence 24125, A
45	19	0.7	1020	US-10-357-930-29216	Sequence 29216, A
46	19	0.7	1024	US-10-105-837-866	Sequence 866, App
47	19	0.7	1099	US-10-115-831-89	Sequence 89, Appl
48	19	0.7	1102	US-10-293-244-335	Sequence 335, App
49	19	0.7	1137	US-10-357-930-24858	Sequence 24858, A
50	19	0.7	1221	US-10-105-837-162	Sequence 162, App
51	19	0.7	1244	US-10-357-930-23272	Sequence 23272, A
52	19	0.7	1244	US-10-357-930-24806	Sequence 24806, A
53	19	0.7	1244	US-10-357-930-28147	Sequence 28147, A
54	19	0.7	1252	US-10-357-930-23154	Sequence 23154, A
55	19	0.7	1252	US-10-357-930-29023	Sequence 29023, A
56	19	0.7	1281	US-10-105-837-972	Sequence 972, App
57	19	0.7	1332	US-10-105-837-196	Sequence 196, App
58	19	0.7	1354	US-10-105-837-308	Sequence 308, App
59	19	0.7	1368	US-10-105-837-871	Sequence 871, App
60	19	0.7	1370	US-10-374-780A-2359	Sequence 2359, App
61	19	0.7	1465	US-10-293-244-646	Sequence 646, App
62	19	0.7	1466	US-10-105-837-914	Sequence 914, App
63	19	0.7	1503	US-10-293-244-415	Sequence 415, App
64	19	0.7	1503	US-10-105-837-588	Sequence 588, App
65	19	0.7	1509	US-10-105-837-722	Sequence 722, App
66	19	0.7	1514	US-10-293-244-748	Sequence 748, App
67	19	0.7	1646	US-10-293-244-576	Sequence 576, App
68	19	0.7	1649	US-10-273-573-4643	Sequence 4643, App
69	19	0.7	1719	US-10-357-930-24297	Sequence 24297, A
70	19	0.7	1719	US-10-357-930-25866	Sequence 25866, A
71	19	0.7	1783	US-10-293-244-297	Sequence 297, App
72	19	0.7	1795	US-10-357-930-25092	Sequence 25092, A
73	19	0.7	1890	US-10-357-930-23084	Sequence 23084, A
74	19	0.7	1930	US-10-293-244-751	Sequence 751, App
75	19	0.7	1986	US-10-357-930-23024	Sequence 23024, A
76	19	0.7	1986	US-10-357-930-28889	Sequence 28889, A
77	19	0.7	1989	US-10-105-837-837	Sequence 837, App
78	19	0.7	2045	US-10-105-837-589	Sequence 589, App
79	19	0.7	2062	US-10-105-837-170	Sequence 170, App
80	19	0.7	2095	US-10-105-837-748	Sequence 748, App
81	19	0.7	2102	US-10-105-837-487	Sequence 487, App
82	19	0.7	2114	US-10-357-930-25185	Sequence 25185, A
83	19	0.7	2210	US-10-105-837-1009	Sequence 1009, App
84	19	0.7	2243	US-10-105-837-1008	Sequence 1008, App
85	19	0.7	2254	US-10-105-837-73	Sequence 73, Appl
86	19	0.7	2261	US-10-115-831-33	Sequence 33, Appl
87	19	0.7	2261	US-10-294-433-33	Sequence 33, Appl
88	19	0.7	2351	US-10-105-837-1007	Sequence 1007, App
89	19	0.7	2376	US-10-357-930-21284	Sequence 21284, A
90	19	0.7	2376	US-10-357-930-21872	Sequence 21872, A
91	19	0.7	2376	US-10-357-930-25228	Sequence 25228, A
92	19	0.7	2376	US-10-357-930-27126	Sequence 27126, A
93	19	0.7	2474	US-10-115-831-121	Sequence 121, App
94	19	0.7	2504	US-10-105-837-778	Sequence 778, App
95	19	0.7	2692	US-10-331-496A-50	Sequence 50, Appl
96	19	0.7	2813	US-10-357-930-30062	Sequence 30062, A
97	19	0.7	2813	US-10-357-930-30123	Sequence 30123, A
98	19	0.7	2874	US-10-105-837-10	Sequence 10, Appl
99	19	0.7	3045	US-10-105-837-9	Sequence 9, Appl

100 19 0.7 3064 6 US-10-105-837-342

Sequence 342, App

ALIGNMENTS

RESULT 1
US-10-293-244-975/c

; Sequence 975, Application US/10293244

; GENERAL INFORMATION:

; APPLICANT: Hyseq, Inc.

; APPLICANT: Tang, Y. Tom et al

; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides

; FILE REFERENCE: 21272-029

; CURRENT APPLICATION NUMBER: US/10/293,244

; CURRENT FILING DATE: 2002-11-12

; PRIOR APPLICATION NUMBER: Not Yet Assigned

; PRIOR FILING DATE: 2001-01-30

; PRIOR APPLICATION NUMBER: 09/728,422

; PRIOR FILING DATE: 2000-11-30

; PRIOR APPLICATION NUMBER: 09/693,325

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 09/663,561

; PRIOR FILING DATE: 2000-09-15

; PRIOR APPLICATION NUMBER: 09/654,936

; PRIOR FILING DATE: 2000-09-01

; PRIOR APPLICATION NUMBER: 09/620,325

; PRIOR FILING DATE: 2000-07-19

; PRIOR APPLICATION NUMBER: 09/598,075

; PRIOR FILING DATE: 2000-06-20

; PRIOR APPLICATION NUMBER: 09/560,875

; PRIOR FILING DATE: 2000-04-27

; PRIOR APPLICATION NUMBER: 09/496,914

; PRIOR FILING DATE: 2000-02-03

; NUMBER OF SEQ ID NOS: 3960

; SOFTWARE: Custom

; SEQ ID NO 975

; LENGTH: 1313

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (547)..(1239)

; US-10-293-244-975

Query Match 1.0%; Score 26; DB 6; Length 1313;
Best Local Similarity 100.0%; Pred. No. 0.00056;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 GTGTCGACCCACGCTCCGGAGTAG 80

DB 1305 GTGTCGACCCACGCTCCGGAGTAG 1280

RESULT 2
PCT-US03-09662-4/c

; Sequence 4, Application PC/TUS0309662

; GENERAL INFORMATION:

; APPLICANT: Bremel, Robert

; APPLICANT: Fakle, Kurt

; APPLICANT: Imboden, Michael

; TITLE OF INVENTION: Antibody Libraries

; FILE REFERENCE: GALA-07772

; CURRENT APPLICATION NUMBER: PCT/US03/09662

; CURRENT FILING DATE: 2003-06-11

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 4

; LENGTH: 7476

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic

PCT-US03-09662-4

Query Match 0.8%; Score 23; DB 1; Length 7476;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATGTTA 23

DB 4495 AATCTTTATTTATCGATGTTA 4473

RESULT 3
PCT-US03-09662-3/c

; Sequence 3, Application PC/TUS0309662

; GENERAL INFORMATION:

; APPLICANT: Bremel, Robert

; APPLICANT: Fakle, Kurt

; APPLICANT: Imboden, Michael

; TITLE OF INVENTION: Antibody Libraries

; FILE REFERENCE: GALA-07772

; CURRENT APPLICATION NUMBER: PCT/US03/09662

; CURRENT FILING DATE: 2003-06-11

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 3

; LENGTH: 7494

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic

; PCT-US03-09662-3

Query Match 0.8%; Score 23; DB 1; Length 7494;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATGTTA 23

DB 4513 AATCTTTATTTATCGATGTTA 4491

RESULT 4
PCT-US03-09662-1/c

; Sequence 1, Application PC/TUS0309662

; GENERAL INFORMATION:

; APPLICANT: Bremel, Robert

; APPLICANT: Fakle, Kurt

; APPLICANT: Imboden, Michael

; TITLE OF INVENTION: Antibody Libraries

; FILE REFERENCE: GALA-07772

; CURRENT APPLICATION NUMBER: PCT/US03/09662

; CURRENT FILING DATE: 2003-06-11

; NUMBER OF SEQ ID NOS: 4

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 1

; LENGTH: 7617

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Synthetic

; PCT-US03-09662-1

Query Match 0.8%; Score 23; DB 1; Length 7617;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATGTTA 23

DB 4625 AATCTTTATTTATCGATGTTA 4603

RESULT 5
PCT-US03-09662-2/c

Sequence 2, Application PC/TUS0309662
GENERAL INFORMATION:
APPLICANT: Bremel, Robert
APPLICANT: Eakle, Kurt
APPLICANT: Imboden, Michael
TITLE OF INVENTION: Antibody Libraries
FILE REFERENCE: GALA-07772
CURRENT APPLICATION NUMBER: PCT/US03/09662
CURRENT FILING DATE: 2003-06-11
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2
LENGTH: 7626
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Synthetic
PCT-US03-09662-2

Query Match 0.8%; Score 23; DB 1; Length 7626;
Best Local Similarity 100.0%; Pred. No. 0.03;
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTTATTCGAGTTA 23
DB 4634 AATCTTTTATTCGAGTTA 4612

RESULT 6
US-10-293-244-459

Sequence 459, Application US/10293244
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
APPLICANT: Tang, Y. Tom et al
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-029
CURRENT APPLICATION NUMBER: US/10/293,244
CURRENT FILING DATE: 2002-11-12
PRIOR APPLICATION NUMBER: Not Yet Assigned
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 09/728,422
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 09/693,325
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/663,561
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/654,936
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 09/620,325
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/598,075
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 3960
SOFTWARE: Custom
SEQ ID NO 459
LENGTH: 906
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (49)..(756)
US-10-293-244-459

Query Match 0.8%; Score 21; DB 6; Length 906;
Best Local Similarity 100.0%; Pred. No. 0.37;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 GTGTGACCCACGCGTCCGGG 75
|||||

DB 17 GTGTGACCCACGCGTCCGGG 37

RESULT 7
US-10-357-930-28772/C

Sequence 22912, Application US/10357930
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Endege, Wilson
APPLICANT: Monahan, John
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
FILE REFERENCE: MRI-007BCN
CURRENT APPLICATION NUMBER: US/10/357,930
CURRENT FILING DATE: 2003-02-04
PRIOR APPLICATION NUMBER: 09/785,276
PRIOR FILING DATE: 2003-02-16
PRIOR APPLICATION NUMBER: 60/183,319
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 60/189,862
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: 60/207,454
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/211,314
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/219,007
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/255,281
PRIOR FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 62232
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 22912
LENGTH: 1026
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 1, 77, 1025, 1026
OTHER INFORMATION: n = A,T,C or G
US-10-357-930-22912

Query Match 0.8%; Score 21; DB 6; Length 1026;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGTGACCCACGCGTCCGGG 77
DB 1024 GTGTGACCCACGCGTCCGGG 1004

RESULT 8
US-10-357-930-28772/C

Sequence 28772, Application US/10357930
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Endege, Wilson
APPLICANT: Monahan, John
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
FILE REFERENCE: MRI-007BCN
CURRENT APPLICATION NUMBER: US/10/357,930
CURRENT FILING DATE: 2003-02-04
PRIOR APPLICATION NUMBER: 09/785,276
PRIOR FILING DATE: 2003-02-16
PRIOR APPLICATION NUMBER: 60/183,319
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 60/189,862
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: 60/207,454
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/211,314

```

; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28772
; LENGTH: 1026
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 77, 1025, 1026
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-28772
```

```
Query Match          0.8%; Score 21; DB 6; Length 1026;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTGACCCACGCGTCCGGGAG 77
Db 1024 GTGACCCACGCGTCCGGGAG 1004
```

```

RESULT 9
US-10-105-837-33
; Sequence 33, Application US/10105837
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing A.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: 784CIP2BDIVA
; CURRENT APPLICATION NUMBER: US/10/105,837
; CURRENT FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_FL_genes Version 1.0
; SEQ ID NO 33
; LENGTH: 1181
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (173)..(874)
US-10-105-837-33
```

```
Query Match          0.8%; Score 21; DB 6; Length 1181;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTGACCCACGCGTCCGGGAG 77
Db 59 GTGACCCACGCGTCCGGGAG 79
```

```

RESULT 10
US-10-293-244-765
; Sequence 765, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
US-10-105-837-664
```

```

; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 765
; LENGTH: 1538
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (237)..(1451)
US-10-293-244-765
```

```
Query Match          0.8%; Score 21; DB 6; Length 1538;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTGACCCACGCGTCCGGGAG 77
Db 50 GTGACCCACGCGTCCGGGAG 70
```

```

RESULT 11
US-10-105-837-664
; Sequence 664, Application US/10105837
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing A.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: 784CIP2BDIVA
; CURRENT APPLICATION NUMBER: US/10/105,837
; CURRENT FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_FL_genes Version 1.0
; SEQ ID NO 664
; LENGTH: 1572
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (330)..(704)
US-10-105-837-664
```

```
Query Match          0.8%; Score 21; DB 6; Length 1572;
```

Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 57 GTGACCCACGCGTCGGGAG 77
Db 280 GTGACCCACGCGTCGGGAG 300

RESULT 12

US-10-357-930-27697
; Sequence 27697, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27697
; LENGTH: 2028
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1804, 1805, 1806, 1807, 1808, 1809, 1810, 1811, 1812, 1813,
; LOCATION: 1814, 1815, 1816, 1817, 1818, 1819, 1820, 1821, 1822, 1823,
; LOCATION: 1824, 1825, 1826, 1827, 1828, 1829, 1830, 1831, 1832, 1833,
; LOCATION: 1834, 1835, 1836, 1837, 1838, 1839, 1840, 1841, 1842
; OTHER INFORMATION: n = A,T,C or G
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1843, 1844, 1845, 1846, 1847, 1848, 1849, 1850, 1851, 1852,
; LOCATION: 1853, 1854, 1855, 1856, 1857, 1858, 1859, 1860, 1861, 1862,
; LOCATION: 2022, 2023, 2024, 2025, 2026, 2027, 2028
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-27697

Query Match 0.8%; Score 21; DB 6; Length 2028;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCGGGAG 77
Db 7 GTGACCCACGCGTCGGGAG 27

RESULT 13

US-09-663-481-2
; Sequence 2, Application US/09663481
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc.
; APPLICANT: Fidock, Mark D.
; TITLE OF INVENTION: PHOSPHODIESTERASE ENZYMES
; FILE REFERENCE: PC10350AGPR

; CURRENT APPLICATION NUMBER: US/09/663,481
; CURRENT FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/177,326
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: UK 9922125.1
; PRIOR FILING DATE: 1999-09-17
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 3091
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-663-481-2

Query Match 0.8%; Score 21; DB 5; Length 3091;
Best Local Similarity 100.0%; Pred. No. 0.39;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCGGGAG 77
Db 1 GTGACCCACGCGTCGGGAG 21

RESULT 14

US-10-229-541A-54
; Sequence 54, Application US/10229541A
; GENERAL INFORMATION:
; APPLICANT: McCourt, Peter
; APPLICANT: Ghaesemian, Majid
; APPLICANT: Cutler, Sean
; APPLICANT: Bonetta, Dario
; TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants
; FILE REFERENCE: 22542-007CIP2
; CURRENT APPLICATION NUMBER: US/10/229,541A
; CURRENT FILING DATE: 2002-08-27
; PRIOR APPLICATION NUMBER: 10/160,764
; PRIOR FILING DATE: 2002-05-31
; PRIOR APPLICATION NUMBER: 60/294,766
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/348,909
; PRIOR FILING DATE: 2001-10-22
; PRIOR APPLICATION NUMBER: 10/210,760
; PRIOR FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: 60/309,396
; PRIOR FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/337,084
; PRIOR FILING DATE: 2001-12-04
; PRIOR APPLICATION NUMBER: 09/191,687
; PRIOR FILING DATE: 1998-11-13
; NUMBER OF SEQ ID NOS: 176
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 54
; LENGTH: 5511
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Plasmid
; OTHER INFORMATION: pBl121-35S-Ant1-AtPTB
US-10-229-541A-54

Query Match 0.8%; Score 21; DB 6; Length 5511;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCGGGAG 77
Db 3349 GTGACCCACGCGTCGGGAG 3369

RESULT 15

US-10-229-541A-79
; Sequence 79, Application US/10229541A
; GENERAL INFORMATION:

```
/ APPLICANT: McCourt, Peter
/ APPLICANT: Ghasssemlan, Majid
/ APPLICANT: Cutler, Sean
/ APPLICANT: Bonetta, Dario
/ TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants
/ FILE REFERENCE: 22542-007CIP2
/ CURRENT APPLICATION NUMBER: US/10/229,541A
/ CURRENT FILING DATE: 2002-08-27
/ PRIOR APPLICATION NUMBER: 10/160,764
/ PRIOR FILING DATE: 2002-05-31
/ PRIOR APPLICATION NUMBER: 60/294,766
/ PRIOR FILING DATE: 2001-05-31
/ PRIOR APPLICATION NUMBER: 60/348,909
/ PRIOR FILING DATE: 2001-10-22
/ PRIOR APPLICATION NUMBER: 10/210,760
/ PRIOR FILING DATE: 2002-08-01
/ PRIOR APPLICATION NUMBER: 60/309,396
/ PRIOR FILING DATE: 2001-08-01
/ PRIOR APPLICATION NUMBER: 60/337,084
/ PRIOR FILING DATE: 2001-12-04
/ PRIOR APPLICATION NUMBER: 09/191,687
/ PRIOR FILING DATE: 1998-11-13
/ NUMBER OF SEQ ID NOS: 176
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 79
/ LENGTH: 5517
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: DN90AFTB
US-10-229-541A-79
```

```
Query Match      0.8%; Score 21; DB 6; Length 5517;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      57  GTGACCCACGCGTCGGGAG 77
Db      3349  GTGACCCACGCGTCGGGAG 3369
```

```
RESULT 16
US-10-229-541A-55
/ Sequence 55, Application US/10229541A
/ GENERAL INFORMATION:
/ APPLICANT: McCourt, Peter
/ APPLICANT: Ghasssemlan, Majid
/ APPLICANT: Cutler, Sean
/ APPLICANT: Bonetta, Dario
/ TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants
/ FILE REFERENCE: 22542-007CIP2
/ CURRENT APPLICATION NUMBER: US/10/229,541A
/ CURRENT FILING DATE: 2002-08-27
/ PRIOR APPLICATION NUMBER: 10/160,764
/ PRIOR FILING DATE: 2002-05-31
/ PRIOR APPLICATION NUMBER: 60/294,766
/ PRIOR FILING DATE: 2001-05-31
/ PRIOR APPLICATION NUMBER: 60/348,909
/ PRIOR FILING DATE: 2001-10-22
/ PRIOR APPLICATION NUMBER: 10/210,760
/ PRIOR FILING DATE: 2002-08-01
/ PRIOR APPLICATION NUMBER: 60/309,396
/ PRIOR FILING DATE: 2001-08-01
/ PRIOR APPLICATION NUMBER: 60/337,084
/ PRIOR FILING DATE: 2001-12-04
/ PRIOR APPLICATION NUMBER: 09/191,687
/ NUMBER OF SEQ ID NOS: 176
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 55
/ LENGTH: 5635
/ TYPE: DNA
```

```
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Plasmid
/ OTHER INFORMATION: pBI12-RD29AP-Anti-AcFTB
US-10-229-541A-55
```

```
Query Match      0.8%; Score 21; DB 6; Length 5635;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      57  GTGACCCACGCGTCGGGAG 77
Db      3473  GTGACCCACGCGTCGGGAG 3493
```

```
RESULT 17
US-10-229-541A-56
/ Sequence 56, Application US/10229541A
/ GENERAL INFORMATION:
/ APPLICANT: McCourt, Peter
/ APPLICANT: Ghasssemlan, Majid
/ APPLICANT: Cutler, Sean
/ APPLICANT: Bonetta, Dario
/ TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants
/ FILE REFERENCE: 22542-007CIP2
/ CURRENT APPLICATION NUMBER: US/10/229,541A
/ CURRENT FILING DATE: 2002-08-27
/ PRIOR APPLICATION NUMBER: 10/160,764
/ PRIOR FILING DATE: 2002-05-31
/ PRIOR APPLICATION NUMBER: 60/294,766
/ PRIOR FILING DATE: 2001-05-31
/ PRIOR APPLICATION NUMBER: 60/348,909
/ PRIOR FILING DATE: 2001-10-22
/ PRIOR APPLICATION NUMBER: 10/210,760
/ PRIOR FILING DATE: 2002-08-01
/ PRIOR APPLICATION NUMBER: 60/309,396
/ PRIOR FILING DATE: 2001-08-01
/ PRIOR APPLICATION NUMBER: 60/337,084
/ PRIOR FILING DATE: 2001-12-04
/ PRIOR APPLICATION NUMBER: 09/191,687
/ PRIOR FILING DATE: 1998-11-13
/ NUMBER OF SEQ ID NOS: 176
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 56
/ LENGTH: 6299
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: Plasmid
US-10-229-541A-56
```

```
Query Match      0.8%; Score 21; DB 6; Length 6299;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      57  GTGACCCACGCGTCGGGAG 77
Db      4883  GTGACCCACGCGTCGGGAG 4903
```

```
RESULT 18
US-10-229-541A-56/c
/ Sequence 56, Application US/10229541A
/ GENERAL INFORMATION:
/ APPLICANT: McCourt, Peter
/ APPLICANT: Ghasssemlan, Majid
/ APPLICANT: Cutler, Sean
/ APPLICANT: Bonetta, Dario
/ TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants
/ FILE REFERENCE: 22542-007CIP2
/ CURRENT APPLICATION NUMBER: US/10/229,541A
/ CURRENT FILING DATE: 2002-08-27
```

PRIOR APPLICATION NUMBER: 10/160,764
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: 60/294,766
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/348,909
PRIOR FILING DATE: 2001-10-22
PRIOR APPLICATION NUMBER: 10/210,760
PRIOR FILING DATE: 2002-08-01
PRIOR APPLICATION NUMBER: 60/309,396
PRIOR FILING DATE: 2001-08-01
PRIOR APPLICATION NUMBER: 60/337,084
PRIOR FILING DATE: 2001-12-04
PRIOR APPLICATION NUMBER: 09/191,687
PRIOR FILING DATE: 1998-11-13
NUMBER OF SEQ ID NOS: 176
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 56
LENGTH: 6299
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Plasmid
US-10-229-541A-56

Query Match 0.8%; Score 21; DB 6; Length 6299;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCGGGAG 77
DB 3819 GTGACCCACGCGTCGGGAG 3799

RESULT 19
US-10-229-541A-57
Sequence 57, Application US/10229541A
GENERAL INFORMATION:
APPLICANT: McCourt, Peter
APPLICANT: Ghassemlan, Majid
APPLICANT: Cutler, Sean
APPLICANT: Bonetta, Dario
TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants
FILE REFERENCE: 22542-007CIP2
CURRENT APPLICATION NUMBER: US/10/229,541A
CURRENT FILING DATE: 2002-08-27
PRIOR APPLICATION NUMBER: 10/160,764
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: 60/294,766
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/348,909
PRIOR FILING DATE: 2001-10-22
PRIOR APPLICATION NUMBER: 10/210,760
PRIOR FILING DATE: 2002-08-01
PRIOR APPLICATION NUMBER: 60/309,396
PRIOR FILING DATE: 2001-08-01
PRIOR APPLICATION NUMBER: 60/337,084
PRIOR FILING DATE: 2001-12-04
PRIOR APPLICATION NUMBER: 09/191,687
PRIOR FILING DATE: 1998-11-13
NUMBER OF SEQ ID NOS: 176
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 57
LENGTH: 6423
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Plasmid
US-10-229-541A-57

Query Match 0.8%; Score 21; DB 6; Length 6423;
Best Local Similarity 100.0%; Pred. No. 0.41;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 57 GTGACCCACGCGTCGGGAG 77
DB 5007 GTGACCCACGCGTCGGGAG 5027

RESULT 20
US-10-229-541A-57/c
Sequence 57, Application US/10229541A
GENERAL INFORMATION:
APPLICANT: McCourt, Peter
APPLICANT: Ghassemlan, Majid
APPLICANT: Cutler, Sean
APPLICANT: Bonetta, Dario
TITLE OF INVENTION: Stress Tolerance and Delayed Senescence in Plants
FILE REFERENCE: 22542-007CIP2
CURRENT APPLICATION NUMBER: US/10/229,541A
CURRENT FILING DATE: 2002-08-27
PRIOR APPLICATION NUMBER: 10/160,764
PRIOR FILING DATE: 2002-05-31
PRIOR APPLICATION NUMBER: 60/294,766
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/348,909
PRIOR FILING DATE: 2001-10-22
PRIOR APPLICATION NUMBER: 10/210,760
PRIOR FILING DATE: 2002-08-01
PRIOR APPLICATION NUMBER: 60/337,084
PRIOR FILING DATE: 2001-12-04
PRIOR APPLICATION NUMBER: 09/191,687
PRIOR FILING DATE: 1998-11-13
NUMBER OF SEQ ID NOS: 176
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 57
LENGTH: 6423
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Plasmid
US-10-229-541A-57

Query Match 0.8%; Score 21; DB 6; Length 6423;
Best Local Similarity 100.0%; Pred. No. 0.41;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCGGGAG 77
DB 3943 GTGACCCACGCGTCGGGAG 3923

RESULT 21
US-10-357-930-50683
Sequence 50683, Application US/10357930
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Bodege, Wilson
APPLICANT: Monahan, John
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF HUMAN PROSTATE CANCER
FILE REFERENCE: MRI-007BCN
CURRENT APPLICATION NUMBER: US/10/357,930
CURRENT FILING DATE: 2003-02-04
PRIOR APPLICATION NUMBER: 09/785,276
PRIOR FILING DATE: 2003-02-16
PRIOR APPLICATION NUMBER: 60/183,319
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 60/189,862
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: 60/207,454

```

; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 50683
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 24
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-50683
```

```

Query Match          0.7%; Score 20; DB 6; Length 598;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 744 TAGATGAGATTAATTAATCTCT 763
DB 269 TAGATGAGATTAATTAATCTCT 288
```

```

RESULT 22
US-10-105-837-562
; Sequence 562, Application US/10105837
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing A.
; APPLICANT: Dimaac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: 784CIP2BDIVA
; CURRENT APPLICATION NUMBER: US/10/105,837
; CURRENT FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: 09/488,725
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_FL_genes Version 1.0
; SEQ ID NO 562
; LENGTH: 1168
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (148)..(615)
US-10-105-837-562
```

```

Query Match          0.7%; Score 20; DB 6; Length 1168;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTGACCCACGCGTCCGGGA 76
DB 25 GTGACCCACGCGTCCGGGA 44
```

```

RESULT 23
US-10-391-363A-59
; Sequence 59, Application US/10391363A
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology LLC
```

```

; APPLICANT: Valentin, Henry E
; APPLICANT: Venkatesh, Tyamagondlu V
; APPLICANT: Karunanandaa, Balasubramini
; TITLE OF INVENTION: Homoglutathione S-transferase ("hPT") Nucleic Acids and Polypeptide
; FILE REFERENCE: REN-02-052 MON-52933
; CURRENT APPLICATION NUMBER: US/10/391,363A
; CURRENT FILING DATE: 2003-03-18
; PRIOR APPLICATION NUMBER: 60/365,202
; PRIOR FILING DATE: 2002-03-19
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 59
; LENGTH: 1423
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-10-391-363A-59
```

```

Query Match          0.7%; Score 20; DB 6; Length 1423;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1372 TACAAAGACCTTCTCTGATG 1391
DB 926 TACAAAGACCTTCTCTGATG 945
```

```

RESULT 24
US-10-115-831-30
; Sequence 30, Application US/10115831
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Dimaac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: 792CIP2ADIV
; CURRENT APPLICATION NUMBER: US/10/115,831
; CURRENT FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: 09/667,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: pc_FL_genes Version 2.0
; SEQ ID NO 30
; LENGTH: 1498
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (162)..(815)
US-10-115-831-30
```

```

Query Match          0.7%; Score 20; DB 6; Length 1498;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTGACCCACGCGTCCGGGA 76
DB 23 GTGACCCACGCGTCCGGGA 42
```

```

RESULT 25
US-10-115-831-102
; Sequence 102, Application US/10115831
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
```

```
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and
FILE REFERENCE: 792CIP2ADIV
CURRENT APPLICATION NUMBER: US/10/115,831
PRIOR FILING DATE: 2002-04-02
PRIOR APPLICATION NUMBER: 09/667,298
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 09/577,408
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 178
SOFTWARE: pt_FL_genes Version 2.0
SEQ ID NO 102
LENGTH: 1522
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (663)..(1238)
US-10-115-831-102
```

```
Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 1522;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTCGACCCACGCGTCGCGGA 76
Db 26 GTCGACCCACGCGTCGCGGA 45
```

```
RESULT 26
US-10-293-244-480
Sequence 480, Application US/10293244
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
APPLICANT: Tang, Y. Tom et al
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-029
CURRENT APPLICATION NUMBER: US/10/293,244
CURRENT FILING DATE: 2002-11-12
PRIOR APPLICATION NUMBER: Not Yet Assigned
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 09/728,422
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 09/693,325
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/663,561
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/654,936
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 09/620,325
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/598,075
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 3960
SOFTWARE: Custom
SEQ ID NO 480
LENGTH: 1906
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (67)..(1740)
US-10-293-244-480
```

```
Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 1906;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTCGACCCACGCGTCGCGGA 76
Db 14 GTCGACCCACGCGTCGCGGA 33
```

```
RESULT 27
US-10-105-837-7
Sequence 7, Application US/10105837
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhao, Qing A.
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and
FILE REFERENCE: 784CIP2BDIVA
CURRENT APPLICATION NUMBER: US/10/105,837
CURRENT FILING DATE: 2002-03-21
PRIOR APPLICATION NUMBER: 09/620,312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 7
LENGTH: 1920
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (120)..(1394)
US-10-105-837-7
```

```
Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 1920;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTCGACCCACGCGTCGCGGA 76
Db 21 GTCGACCCACGCGTCGCGGA 40
```

```
RESULT 28
US-10-105-837-1046
Sequence 1046, Application US/10105837
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhao, Qing A.
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and
FILE REFERENCE: 784CIP2BDIVA
CURRENT APPLICATION NUMBER: US/10/105,837
CURRENT FILING DATE: 2002-03-21
PRIOR APPLICATION NUMBER: 09/620,312
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 1046
LENGTH: 2225
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
```

```
/ LOCATION: (260)...(1906)
US-10-105-837-1046

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 2225;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTCGACCCACGCGTCCGGGA 76
Db 35 GTCGACCCACGCGTCCGGGA 54

RESULT 29
US-10-357-930-22213
Sequence 22213, Application US/10357930
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Endege, Wilson
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
FILE REFERENCE: MRI-007BCN
CURRENT APPLICATION NUMBER: US/10/357,930
CURRENT FILING DATE: 2003-02-04
PRIOR APPLICATION NUMBER: 09/785,276
PRIOR FILING DATE: 2003-02-16
PRIOR APPLICATION NUMBER: 60/183,319
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 60/189,862
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: 60/207,454
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/211,314
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/219,007
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/255,281
PRIOR FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 62232
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 22213
LENGTH: 2946
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 1, 2945, 2946
OTHER INFORMATION: n = A,T,C or G
US-10-357-930-22213

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 2946;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTCGACCCACGCGTCCGGGA 76
Db 9 GTCGACCCACGCGTCCGGGA 28

RESULT 30
US-10-357-930-23051
Sequence 23051, Application US/10357930
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Endege, Wilson
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
FILE REFERENCE: MRI-007BCN
CURRENT APPLICATION NUMBER: US/10/357,930
CURRENT FILING DATE: 2003-02-04
```

```
/ PRIOR APPLICATION NUMBER: 09/785,276
/ PRIOR FILING DATE: 2003-02-16
/ PRIOR APPLICATION NUMBER: 60/183,319
/ PRIOR FILING DATE: 2000-02-17
/ PRIOR APPLICATION NUMBER: 60/189,862
/ PRIOR FILING DATE: 2000-03-16
/ PRIOR APPLICATION NUMBER: 60/207,454
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: 60/211,314
/ PRIOR FILING DATE: 2000-06-09
/ PRIOR APPLICATION NUMBER: 60/219,007
/ PRIOR FILING DATE: 2000-07-18
/ PRIOR APPLICATION NUMBER: 60/255,281
/ PRIOR FILING DATE: 2000-12-13
/ NUMBER OF SEQ ID NOS: 62232
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 23051
/ LENGTH: 2946
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: 1, 2945, 2946
/ OTHER INFORMATION: n = A,T,C or G
US-10-357-930-23051

Query Match
Best Local Similarity 100.0%; Score 20; DB 6; Length 2946;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTCGACCCACGCGTCCGGGA 76
Db 9 GTCGACCCACGCGTCCGGGA 28

RESULT 31
US-10-357-930-28070
Sequence 28070, Application US/10357930
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Endege, Wilson
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
FILE REFERENCE: MRI-007BCN
CURRENT APPLICATION NUMBER: US/10/357,930
CURRENT FILING DATE: 2003-02-04
PRIOR APPLICATION NUMBER: 09/785,276
PRIOR FILING DATE: 2003-02-16
PRIOR APPLICATION NUMBER: 60/183,319
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 60/189,862
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: 60/207,454
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/211,314
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/219,007
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/255,281
PRIOR FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 62232
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 28070
LENGTH: 2946
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 1, 2945, 2946
OTHER INFORMATION: n = A,T,C or G
US-10-357-930-28070
```


Query Match 0.7%; Score 20; DB 6; Length 2946;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGGCTCCGGGA 76
|||||
Db 9 GTGACCCACGGCTCCGGGA 28

RESULT 32

US-10-357-930-28916
Sequence 28916, Application US/10357930
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Endege, Wilson
APPLICANT: Monahan, John
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
FILE REFERENCE: MRI-0078CN
CURRENT APPLICATION NUMBER: US/10/357,930
CURRENT FILING DATE: 2003-02-04
PRIOR APPLICATION NUMBER: 09/785,276
PRIOR FILING DATE: 2003-02-16
PRIOR APPLICATION NUMBER: 60/183,319
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 60/189,862
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: 60/207,454
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/211,314
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/219,007
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/255,281
PRIOR FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 62232
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 28916
LENGTH: 2946
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 1, 2545, 2946
OTHER INFORMATION: n = A,T,C or G
US-10-357-930-28916

Query Match 0.7%; Score 20; DB 6; Length 2946;
Best Local Similarity 100.0%; Pred. No. 1.4;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGGCTCCGGGA 76
|||||
Db 9 GTGACCCACGGCTCCGGGA 28

RESULT 33

US-09-654-936A-134
Sequence 134, Application US/09654936A
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhou, Ping
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Zhao, Qiong A.
APPLICANT: Xue, Aidong J.
APPLICANT: Wehrman, Tom
APPLICANT: Wang, Jian-Rui
APPLICANT: Dimanac, Radoje T.

TITLE OF INVENTION: Novel Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 787CIP2C
CURRENT APPLICATION NUMBER: US/09/654,936A
CURRENT FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 164
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 134
LENGTH: 3336
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (132)..(1751)
US-09-654-936A-134

Query Match 0.7%; Score 20; DB 5; Length 3336;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 TCGACCCACGGCTCCGGGAG 77
|||||
Db 2 TCGACCCACGGCTCCGGGAG 21

RESULT 34

US-10-293-244-900
Sequence 900, Application US/10293244
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc.
APPLICANT: Tang, Y. Tom et al
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-029
CURRENT APPLICATION NUMBER: US/10/293,244
CURRENT FILING DATE: 2002-11-12
PRIOR APPLICATION NUMBER: Not Yet Assigned
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 09/728,422
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 09/693,325
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/663,561
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/654,936
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 09/620,325
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/598,075
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
PRIOR FILING DATE: 2000-02-03
NUMBER OF SEQ ID NOS: 3960
SOFTWARE: Custom
SEQ ID NO 900
LENGTH: 3336
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (132)..(1751)
US-10-293-244-900

Query Match 0.7%; Score 20; DB 6; Length 3336;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 TCGACCCACGGCTCCGGGAG 77

Db 2 TCGACCACGCGCGCGAG 21

RESULT 35
US-09-947-914-41/C
; Sequence 41, Application US/09947914
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig
; TITLE OF INVENTION: COLLECTION OF SINGLE NUCLEOTIDE POLYMORPHISM (SNPs) LOCATED ON CH
; FILE REFERENCE: CL001298
; CURRENT APPLICATION NUMBER: US/09/947,914
; PRIOR FILING DATE: 2001-09-07
; NUMBER OF SEQ ID NOS: 75
; SEQ ID NO 41
; LENGTH: 13831263
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(13831263)
; OTHER INFORMATION: n = A,T,C or G
US-09-947-914-41

Query Match 0.7%; Score 20; DB 5; Length 13831263;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1962 TAGAGATAGTACACATTT 1981
Db 4127333 TAGAGATAGTACACATTT 4127314

RESULT 36
US-10-273-573-1607/C
; Sequence 1607, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; PRIOR FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 1607
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (362)...(252)
; OTHER INFORMATION: 97% homologous to Oryctolagus cuniculus SNRPn upstream
; OTHER INFORMATION: reading frame protein, accession number AF101043, Smith-Waterman
; OTHER INFORMATION: Score=181.
US-10-273-573-1607

Query Match 0.7%; Score 19; DB 6; Length 433;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCAACGCGTCCGG 75
Db 429 GTGACCAACGCGTCCGG 411

RESULT 37
US-10-085-783A-26572
; Sequence 26572, Application US/10085783A

; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 26572
; LENGTH: 470
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (16)...(16)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-26572

Query Match 0.7%; Score 19; DB 6; Length 470;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2691 GGCTTTTCTTCTAATACC 2709
Db 334 GGCTTTTCTTCTAATACC 352

RESULT 38
US-10-085-783A-39690
; Sequence 39690, Application US/10085783A
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 39690
; LENGTH: 479
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (356)...(356)
; OTHER INFORMATION: n is a, c, g, or t
US-10-085-783A-39690

Query Match 0.7%; Score 19; DB 6; Length 479;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1615 AGAAGAGTTCTGTGAA 1633
Db 251 AGAAGAGTTCTGTGAA 269

RESULT 39
US-10-357-930-21683

```
/ Sequence 21683, Application US/10357930
/ GENERAL INFORMATION:
/ APPLICANT: Schlegel, Robert
/ APPLICANT: Endege, Wilson
/ APPLICANT: Monahan, John
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
/ FILE REFERENCE: MRI-0078CN
/ CURRENT APPLICATION NUMBER: US/10/357,930
/ CURRENT FILING DATE: 2003-02-04
/ PRIOR APPLICATION NUMBER: 09/785,276
/ PRIOR FILING DATE: 2003-02-16
/ PRIOR APPLICATION NUMBER: 60/183,319
/ PRIOR FILING DATE: 2000-02-17
/ PRIOR APPLICATION NUMBER: 60/189,862
/ PRIOR FILING DATE: 2000-03-16
/ PRIOR APPLICATION NUMBER: 60/207,454
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: 60/211,314
/ PRIOR FILING DATE: 2000-06-09
/ PRIOR APPLICATION NUMBER: 60/219,007
/ PRIOR FILING DATE: 2000-07-18
/ PRIOR APPLICATION NUMBER: 60/255,281
/ PRIOR FILING DATE: 2000-12-13
/ SOFTWARE: FastSeq for Windows Version 4.0
/ NUMBER OF SEQ ID NOS: 62232
/ SEQ ID NO 21683
/ LENGTH: 530
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: 527..528, 529, 530
/ OTHER INFORMATION: n = A,T,C or G
US-10-357-930-21683

Query Match          0.7%; Score 19; DB 6; Length 530;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTCGACCCACGCGTCGCGG 75
      |||||||
Db      2 GTCGACCCACGCGTCGCGG 20

RESULT 40
US-10-357-930-27527
/ Sequence 27527, Application US/10357930
/ GENERAL INFORMATION:
/ APPLICANT: Schlegel, Robert
/ APPLICANT: Endege, Wilson
/ APPLICANT: Monahan, John
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
/ FILE REFERENCE: MRI-0078CN
/ CURRENT APPLICATION NUMBER: US/10/357,930
/ CURRENT FILING DATE: 2003-02-04
/ PRIOR APPLICATION NUMBER: 09/785,276
/ PRIOR FILING DATE: 2003-02-16
/ PRIOR APPLICATION NUMBER: 60/183,319
/ PRIOR FILING DATE: 2000-02-17
/ PRIOR APPLICATION NUMBER: 60/189,862
/ PRIOR FILING DATE: 2000-03-16
/ PRIOR APPLICATION NUMBER: 60/207,454
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: 60/211,314
/ PRIOR FILING DATE: 2000-06-09
/ PRIOR APPLICATION NUMBER: 60/219,007
/ PRIOR FILING DATE: 2000-07-18
/ PRIOR APPLICATION NUMBER: 60/255,281
/ PRIOR FILING DATE: 2000-12-13
```

```
/ NUMBER OF SEQ ID NOS: 62232
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 27527
/ LENGTH: 530
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: 527..528, 529, 530
/ OTHER INFORMATION: n = A,T,C or G
US-10-357-930-27527

Query Match          0.7%; Score 19; DB 6; Length 530;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTCGACCCACGCGTCGCGG 75
      |||||||
Db      2 GTCGACCCACGCGTCGCGG 20

RESULT 41
US-10-105-837-920
/ Sequence 920, Application US/10105837
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Dmanac, Radoje T.
/ TITLE OF INVENTION: Novel Nucleic Acids and
/ TITLE OF INVENTION: Polypeptides
/ FILE REFERENCE: 784CIP2BDIVA
/ CURRENT APPLICATION NUMBER: US/10/105,837
/ CURRENT FILING DATE: 2002-03-21
/ PRIOR APPLICATION NUMBER: 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1104
/ SOFTWARE: pc_fl_genes Version 1.0
/ SEQ ID NO 920
/ LENGTH: 571
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (63)..(425)
US-10-105-837-920

Query Match          0.7%; Score 19; DB 6; Length 571;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTCGACCCACGCGTCGCGG 75
      |||||||
Db      31 GTCGACCCACGCGTCGCGG 49

RESULT 42
US-10-273-4669/C
/ Sequence 4669, Application US/10273573
/ GENERAL INFORMATION:
/ APPLICANT: Hyseq, Inc
/ TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
/ FILE REFERENCE: 21272-066
/ CURRENT APPLICATION NUMBER: US/10/273,573
/ CURRENT FILING DATE: 2002-10-18
/ PRIOR APPLICATION NUMBER: 09/522,929
/ PRIOR FILING DATE: 2000-04-18
/ PRIOR APPLICATION NUMBER: 09/770,160
```

```
/ PRIOR FILING DATE: 2001-01-26
/ NUMBER OF SEQ ID NOS: 10994
/ SOFTWARE: Custom
/ SEQ ID NO 4669
/ LENGTH: 786
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: SIMILAR
/ LOCATION: (161)..(27)
/ OTHER INFORMATION: 45% homologous to Homo sapiens type I procollagen, accession
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)..(786)
/ OTHER INFORMATION: n = a,t,c or g
US-10-273-573-4669
```

```
Query Match 0.7%; Score 19; DB 6; Length 786;
Best Local Similarity 100.0%; Pred. No. 5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 58 TCGACCCACGCGTCCGGGA 76
DB 774 TCGACCCACGCGTCCGGGA 756
```

```
RESULT 43
US-10-357-930-23333
/ Sequence 23333, Application US/10357930
/ GENERAL INFORMATION:
/ APPLICANT: Schlegel, Robert
/ APPLICANT: Endege, Wilson
/ APPLICANT: Monahan, John
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
/ FILE REFERENCE: MRI-007BCN
/ CURRENT APPLICATION NUMBER: US/10/357,930
/ CURRENT FILING DATE: 2003-02-04
/ PRIOR APPLICATION NUMBER: 09/785,276
/ PRIOR FILING DATE: 2003-02-16
/ PRIOR APPLICATION NUMBER: 60/183,319
/ PRIOR FILING DATE: 2000-02-17
/ PRIOR APPLICATION NUMBER: 60/189,862
/ PRIOR FILING DATE: 2000-03-16
/ PRIOR APPLICATION NUMBER: 60/207,454
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: 60/211,314
/ PRIOR FILING DATE: 2000-06-09
/ PRIOR APPLICATION NUMBER: 60/219,007
/ PRIOR FILING DATE: 2000-07-18
/ PRIOR APPLICATION NUMBER: 60/255,281
/ PRIOR FILING DATE: 2000-12-13
/ NUMBER OF SEQ ID NOS: 62232
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 23333
/ LENGTH: 1020
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 1017, 1018, 1019, 1020
/ OTHER INFORMATION: n = A,T,C or G
US-10-357-930-23333
```

```
Query Match 0.7%; Score 19; DB 6; Length 1020;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTCGACCCACGCGTCCGGG 75
DB 6 GTCGACCCACGCGTCCGGG 24
```

```
RESULT 44
US-10-357-930-24125
/ Sequence 24125, Application US/10357930
/ GENERAL INFORMATION:
/ APPLICANT: Schlegel, Robert
/ APPLICANT: Endege, Wilson
/ APPLICANT: Monahan, John
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
/ FILE REFERENCE: MRI-007BCN
/ CURRENT APPLICATION NUMBER: US/10/357,930
/ CURRENT FILING DATE: 2003-02-04
/ PRIOR APPLICATION NUMBER: 09/785,276
/ PRIOR FILING DATE: 2003-02-16
/ PRIOR APPLICATION NUMBER: 60/183,319
/ PRIOR FILING DATE: 2000-02-17
/ PRIOR APPLICATION NUMBER: 60/189,862
/ PRIOR FILING DATE: 2000-03-16
/ PRIOR APPLICATION NUMBER: 60/207,454
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: 60/211,314
/ PRIOR FILING DATE: 2000-06-09
/ PRIOR APPLICATION NUMBER: 60/219,007
/ PRIOR FILING DATE: 2000-07-18
/ PRIOR APPLICATION NUMBER: 60/255,281
/ PRIOR FILING DATE: 2000-12-13
/ NUMBER OF SEQ ID NOS: 62232
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 24125
/ LENGTH: 1020
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 1017, 1018, 1019, 1020
/ OTHER INFORMATION: n = A,T,C or G
US-10-357-930-24125
```

```
Query Match 0.7%; Score 19; DB 6; Length 1020;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTCGACCCACGCGTCCGGG 75
DB 6 GTCGACCCACGCGTCCGGG 24
```

```
RESULT 45
US-10-357-930-29216
/ Sequence 29216, Application US/10357930
/ GENERAL INFORMATION:
/ APPLICANT: Schlegel, Robert
/ APPLICANT: Endege, Wilson
/ APPLICANT: Monahan, John
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
/ FILE REFERENCE: MRI-007BCN
/ CURRENT APPLICATION NUMBER: US/10/357,930
/ CURRENT FILING DATE: 2003-02-04
/ PRIOR APPLICATION NUMBER: 09/785,276
/ PRIOR FILING DATE: 2003-02-16
/ PRIOR APPLICATION NUMBER: 60/183,319
/ PRIOR FILING DATE: 2000-02-17
/ PRIOR APPLICATION NUMBER: 60/189,862
/ PRIOR FILING DATE: 2000-03-16
/ PRIOR APPLICATION NUMBER: 60/207,454
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: 60/211,314
/ PRIOR FILING DATE: 2000-06-09
```

```

; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29216
; LENGTH: 1020
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1017, 1018, 1019, 1020
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-29216

Query Match
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCGGG 75
Db 6 GTGACCCACGCGTCGGG 24

RESULT 46
US-10-105-837-866
; Sequence 866, Application US/10105837
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing A.
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: 784CIP2B.DIVA
; CURRENT APPLICATION NUMBER: US/10/105,837
; PRIOR FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 866
; LENGTH: 1024
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (519)..(980)
US-10-105-837-866

Query Match
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCGGG 75
Db 14 GTGACCCACGCGTCGGG 32

RESULT 47
US-10-115-831-89
; Sequence 89, Application US/10115831
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
```

```

; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: 792CIP2B.DIVA
; CURRENT APPLICATION NUMBER: US/10/115,831
; PRIOR FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: 09/667,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 89
; LENGTH: 1099
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (140)..(445)
US-10-115-831-89

Query Match
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCGGG 75
Db 37 GTGACCCACGCGTCGGG 55

RESULT 48
US-10-293-244-335
; Sequence 335, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hysq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO 335
; LENGTH: 1102
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (76)..(408)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1102)
; OTHER INFORMATION: n = a,t,c or g
US-10-293-244-335
```

Query Match 0.7%; Score 19; DB 6; Length 1102;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGG 75
DB 43 GTGACCCACGCGTCCGGG 61

RESULT 49

US-10-357-930-24858/c
; Sequence 24858, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24858
; LENGTH: 1137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1, 2, 3, 4, 1136, 1137
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-24858

Query Match 0.7%; Score 19; DB 6; Length 1137;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGG 75
DB 1135 GTGACCCACGCGTCCGGG 1117

RESULT 50

US-10-105-837-162
; Sequence 162, Application US/10105837
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhao, Qing A.
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BDIVA
; CURRENT APPLICATION NUMBER: US/10/105,837
; PRIOR FILING DATE: 2002-03-21
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19

; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 162
; LENGTH: 1221
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (263) ..(862)
US-10-105-837-162

US-10-105-837-162

Query Match 0.7%; Score 19; DB 6; Length 1221;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGG 75
DB 19 GTGACCCACGCGTCCGGG 37

RESULT 51

US-10-357-930-23272
; Sequence 23272, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23272
; LENGTH: 1244
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-23272

Query Match 0.7%; Score 19; DB 6; Length 1244;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGG 75
DB 3 GTGACCCACGCGTCCGGG 21

RESULT 52

```
US-10-357-930-24806
; Sequence 24806, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24806
; LENGTH: 1244
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-24806

Query Match          0.7%; Score 19; DB 6; Length 1244;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTCGACCCACGCGTCGCGG 75
Db      3 GTCGACCCACGCGTCGCGG 21

RESULT 53
US-10-357-930-29147
; Sequence 29147, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
```

```
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29147
; LENGTH: 1244
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-29147

Query Match          0.7%; Score 19; DB 6; Length 1244;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTCGACCCACGCGTCGCGG 75
Db      3 GTCGACCCACGCGTCGCGG 21

RESULT 54
US-10-357-930-23154/c
; Sequence 23154, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23154
; LENGTH: 1252
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-23154

Query Match          0.7%; Score 19; DB 6; Length 1252;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTCGACCCACGCGTCGCGG 75
Db      1251 GTCGACCCACGCGTCGCGG 1233

RESULT 55
US-10-357-930-29023/c
; Sequence 29023, Application US/10357930
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Schlegel, Robert
/ APPLICANT: Endege, Wilson
/ APPLICANT: Monahan, John
/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
/ TITLE OF INVENTION: HUMAN PROSTATE CANCER
/ FILE REFERENCE: MRI-007BCN
/ CURRENT APPLICATION NUMBER: US/10/357,930
/ CURRENT FILING DATE: 2003-02-04
/ PRIOR APPLICATION NUMBER: 09/785,276
/ PRIOR FILING DATE: 2003-02-16
/ PRIOR APPLICATION NUMBER: 60/183,319
/ PRIOR FILING DATE: 2000-02-17
/ PRIOR APPLICATION NUMBER: 60/189,862
/ PRIOR FILING DATE: 2000-03-16
/ PRIOR APPLICATION NUMBER: 60/207,454
/ PRIOR FILING DATE: 2000-05-25
/ PRIOR APPLICATION NUMBER: 60/211,314
/ PRIOR FILING DATE: 2000-06-09
/ PRIOR APPLICATION NUMBER: 60/219,007
/ PRIOR FILING DATE: 2000-07-18
/ PRIOR APPLICATION NUMBER: 60/255,281
/ PRIOR FILING DATE: 2000-12-13
/ NUMBER OF SEQ ID NOS: 62232
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 29023
/ LENGTH: 1252
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: 1
/ OTHER INFORMATION: n = A,T,C or G
/ US-10-357-930-29023

Query Match          0.7%; Score 19; DB 6; Length 1252;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTGACCCACGCGTCGGG 75
      |||||
DB      1251 GTGACCCACGCGTCGGG 1233

RESULT 56
/ US-10-105-837-972
/ Sequence 972, Application US/10105837
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: Novel Nucleic Acids and
/ TITLE OF INVENTION: Polypeptides
/ FILE REFERENCE: 784CIP2BDIVA
/ CURRENT APPLICATION NUMBER: US/10/105,837
/ CURRENT FILING DATE: 2002-03-21
/ PRIOR APPLICATION NUMBER: 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1104
/ SOFTWARE: pt_FL_genes Version 1.0
/ SEQ ID NO 972
/ LENGTH: 1281
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
```

```
/ LOCATION: (172) .. (1074)
/ US-10-105-837-972

Query Match          0.7%; Score 19; DB 6; Length 1281;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTGACCCACGCGTCGGG 75
      |||||
DB      14 GTGACCCACGCGTCGGG 32

RESULT 57
/ US-10-105-837-196
/ Sequence 196, Application US/10105837
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: Novel Nucleic Acids and
/ TITLE OF INVENTION: Polypeptides
/ FILE REFERENCE: 784CIP2BDIVA
/ CURRENT APPLICATION NUMBER: US/10/105,837
/ CURRENT FILING DATE: 2002-03-21
/ PRIOR APPLICATION NUMBER: 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1104
/ SOFTWARE: pt_FL_genes Version 1.0
/ SEQ ID NO 196
/ LENGTH: 1332
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (171) .. (920)
/ US-10-105-837-196

Query Match          0.7%; Score 19; DB 6; Length 1332;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTGACCCACGCGTCGGG 75
      |||||
DB      40 GTGACCCACGCGTCGGG 58

RESULT 58
/ US-10-105-837-308
/ Sequence 308, Application US/10105837
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: Novel Nucleic Acids and
/ TITLE OF INVENTION: Polypeptides
/ FILE REFERENCE: 784CIP2BDIVA
/ CURRENT APPLICATION NUMBER: US/10/105,837
/ CURRENT FILING DATE: 2002-03-21
/ PRIOR APPLICATION NUMBER: 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1104
```



```
SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 308
; LENGTH: 1354
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (226)..(1251)
US-10-105-837-308
```

```
Query Match
Best Local Similarity 100.0%; Score 19; DB 6; Length 1354;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 57 GTGACCCACGCGTCGGG 75
|||
Db 129 GTGACCCACGCGTCGGG 147
```

```
RESULT 59
US-10-105-837-871
; Sequence 871, Application US/10105837
; GENERAL INFORMATION:
```

```
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhao, Qing A.
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Novel Nucleic Acids and
FILE REFERENCE: 784CIP2BDIVA
CURRENT FILING DATE: 2002-03-21
PRIOR APPLICATION NUMBER: US/10/105,837
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 871
LENGTH: 1368
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (103)..(447)
US-10-105-837-871
```

```
Query Match
Best Local Similarity 100.0%; Score 19; DB 6; Length 1368;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 57 GTGACCCACGCGTCGGG 75
|||
Db 48 GTGACCCACGCGTCGGG 66
```

```
RESULT 60
US-10-374-780A-2359
; Sequence 2359, Application US/10374780A
; GENERAL INFORMATION:
```

```
APPLICANT: Sherman, Bradley K
APPLICANT: Riechmann, Jose Luis
APPLICANT: Jiang, Cai-Zhong
APPLICANT: Heard, Jacqueline E
APPLICANT: Haake, Volker
APPLICANT: Creelman, Robert A
APPLICANT: Ratcliffe, Oliver
APPLICANT: Adam, Luc J
APPLICANT: Reuber, T. Lynne
APPLICANT: Keddie, James
```

```
APPLICANT: Brown, Pierre E
APPLICANT: Pilgrim, Marsha L
APPLICANT: Dubell III, Arnold T
APPLICANT: Pineda, Omeira
```

```
APPLICANT: Yu, Guo-Liang
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES IN PLANTS
FILE REFERENCE: MBI-0047 CIP
CURRENT APPLICATION NUMBER: US/10/374,780A
CURRENT FILING DATE: 2003-02-25
PRIOR APPLICATION NUMBER: 09/837,944
```

```
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: 60/310,847
PRIOR FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: 09/934,455
PRIOR FILING DATE: 2001-08-22
PRIOR APPLICATION NUMBER: 60/336,049
PRIOR FILING DATE: 2001-11-19
PRIOR APPLICATION NUMBER: 60/338,692
PRIOR FILING DATE: 2001-12-11
PRIOR APPLICATION NUMBER: 10/171,468
PRIOR FILING DATE: 2002-06-14
PRIOR APPLICATION NUMBER: 10/225,066
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: 10/225,067
PRIOR FILING DATE: 2002-08-09
PRIOR APPLICATION NUMBER: 10/225,068
PRIOR FILING DATE: 2002-08-09
NUMBER OF SEQ ID NOS: 2906
SOFTWARE: PatentIn version 3.2
SEQ ID NO 2359
LENGTH: 1370
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: G569
US-10-374-780A-2359
```

```
Query Match
Best Local Similarity 100.0%; Score 19; DB 6; Length 1370;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 57 GTGACCCACGCGTCGGG 75
|||
Db 1 GTGACCCACGCGTCGGG 19
```

```
RESULT 61
US-10-293-244-646
; Sequence 646, Application US/10293244
; GENERAL INFORMATION:
```

```
APPLICANT: Hyseq, Inc.
APPLICANT: Tang, Y. Tom et al
TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
FILE REFERENCE: 21272-029
CURRENT APPLICATION NUMBER: US/10/293,244
CURRENT FILING DATE: 2002-11-12
PRIOR APPLICATION NUMBER: Not Yet Assigned
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: 09/728,422
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: 09/693,325
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 09/663,561
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: 09/654,936
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: 09/620,325
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/598,075
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: 09/496,914
```

```
/ PRIOR FILING DATE: 2000-02-03
/ NUMBER OF SEQ ID NOS: 3960
/ SOFTWARE: Custom
/ SEQ ID NO 646
/ LENGTH: 1465
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (251)..(931)
US-10-293-244-646
```

```
Query Match      0.7%; Score 19; DB 6; Length 1465;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTGACCCACGCGTCCGGG 75
DB 43 GTGACCCACGCGTCCGGG 61
```

```
RESULT 62
US-10-105-837-914
/ Sequence 914, Application US/10105837
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: Novel Nucleic Acids and
/ FILE REFERENCE: 784CIP2B.DIVA
/ CURRENT APPLICATION NUMBER: US/10/105,837
/ CURRENT FILING DATE: 2002-03-21
/ PRIOR APPLICATION NUMBER: 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1104
/ SOFTWARE: pt_FL_genes Version 1.0
/ SEQ ID NO 914
/ LENGTH: 1466
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (779)..(1213)
/ NAME/KEY: misc feature
/ LOCATION: (1)..(1466)
/ OTHER INFORMATION: n = a,t,c or g
US-10-105-837-914
```

```
Query Match      0.7%; Score 19; DB 6; Length 1466;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTGACCCACGCGTCCGGG 75
DB 7 GTGACCCACGCGTCCGGG 25
```

```
RESULT 63
US-10-293-244-415
/ Sequence 415, Application US/10293244
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Hysbeg, Inc.
/ APPLICANT: Tang, Y. Tom et al
/ TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
/ FILE REFERENCE: 21272-029
```

```
/ CURRENT APPLICATION NUMBER: US/10/293,244
/ CURRENT FILING DATE: 2002-11-12
/ PRIOR APPLICATION NUMBER: Not Yet Assigned
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: 09/728,422
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: 09/693,325
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 09/663,561
/ PRIOR FILING DATE: 2000-09-15
/ PRIOR APPLICATION NUMBER: 09/654,936
/ PRIOR FILING DATE: 2000-09-01
/ PRIOR APPLICATION NUMBER: 09/620,325
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/598,075
/ PRIOR FILING DATE: 2000-06-20
/ PRIOR APPLICATION NUMBER: 09/560,875
/ PRIOR FILING DATE: 2000-04-27
/ PRIOR APPLICATION NUMBER: 09/496,914
/ PRIOR FILING DATE: 2000-02-03
/ NUMBER OF SEQ ID NOS: 3960
/ SOFTWARE: Custom
/ SEQ ID NO 415
/ LENGTH: 1503
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (125)..(1474)
US-10-293-244-415
```

```
Query Match      0.7%; Score 19; DB 6; Length 1503;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 57 GTGACCCACGCGTCCGGG 75
DB 36 GTGACCCACGCGTCCGGG 54
```

```
RESULT 64
US-10-105-837-588
/ Sequence 588, Application US/10105837
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: Novel Nucleic Acids and
/ FILE REFERENCE: 784CIP2B.DIVA
/ CURRENT APPLICATION NUMBER: US/10/105,837
/ CURRENT FILING DATE: 2002-03-21
/ PRIOR APPLICATION NUMBER: 09/620,312
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1104
/ SOFTWARE: pt_FL_genes Version 1.0
/ SEQ ID NO 588
/ LENGTH: 1503
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (575)..(1000)
US-10-105-837-588
```

```
Query Match      0.7%; Score 19; DB 6; Length 1503;
Best Local Similarity 100.0%; Pred. No. 5.2;
```

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCGGG 75
|||||
DB 40 GTGACCCACGCGTCGGG 58

RESULT 65

US-10-105-837-722
; Sequence 722, Application US/10105837

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhao, Qing A.

APPLICANT: Drmanac, Radoje T.

TITLE OF INVENTION: Novel Nucleic Acids and

FILE REFERENCE: 784CIP2BDIVA

CURRENT FILING DATE: 2002-03-21

PRIOR APPLICATION NUMBER: 09/620,312

PRIOR FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/552,317

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/488,725

PRIOR FILING DATE: 2000-01-21

NUMBER OF SEQ ID NOS: 1104

SOFTWARE: pt_FL_genes Version 1.0

SEQ ID NO 722

LENGTH: 1509

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (557)..(1312)

FEATURE:

NAME/KEY: misc.feature

LOCATION: (1)..(1509)

OTHER INFORMATION: n = a,t,c or g

US-10-105-837-722

Query Match

Best Local Similarity 0.7%; Score 19; DB 6; Length 1509;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCGGG 75
|||||
DB 55 GTGACCCACGCGTCGGG 73

RESULT 66

US-10-293-244-748

; Sequence 748, Application US/10293244

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.

APPLICANT: Tang, Y. Tom et al

TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides

FILE REFERENCE: 21272-029

CURRENT FILING DATE: 2002-11-12

PRIOR APPLICATION NUMBER: Not Yet Assigned

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: 09/728,422

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: 09/693,325

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 09/663,561

PRIOR FILING DATE: 2000-09-15

PRIOR APPLICATION NUMBER: 09/654,936

PRIOR FILING DATE: 2000-09-01

PRIOR APPLICATION NUMBER: 09/620,325

PRIOR FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/598,075

PRIOR FILING DATE: 2000-06-20

PRIOR APPLICATION NUMBER: 09/560,875

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: 09/496,914

PRIOR FILING DATE: 2000-02-03

NUMBER OF SEQ ID NOS: 3960

SOFTWARE: Custom

SEQ ID NO 748

LENGTH: 1514

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (103)..(1374)

US-10-293-244-748

RESULT 67

US-10-293-244-576

; Sequence 576, Application US/10293244

GENERAL INFORMATION:

APPLICANT: Hyseq, Inc.

APPLICANT: Tang, Y. Tom et al

TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides

FILE REFERENCE: 21272-029

CURRENT FILING DATE: 2002-11-12

PRIOR APPLICATION NUMBER: Not Yet Assigned

PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: 09/728,422

PRIOR FILING DATE: 2000-11-30

PRIOR APPLICATION NUMBER: 09/693,325

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 09/663,561

PRIOR FILING DATE: 2000-09-15

PRIOR APPLICATION NUMBER: 09/654,936

PRIOR FILING DATE: 2000-09-01

PRIOR APPLICATION NUMBER: 09/620,325

PRIOR FILING DATE: 2000-07-19

PRIOR APPLICATION NUMBER: 09/598,075

PRIOR FILING DATE: 2000-06-20

PRIOR APPLICATION NUMBER: 09/560,875

PRIOR FILING DATE: 2000-04-27

PRIOR APPLICATION NUMBER: 09/496,914

PRIOR FILING DATE: 2000-02-03

NUMBER OF SEQ ID NOS: 3960

SOFTWARE: Custom

SEQ ID NO 576

LENGTH: 1646

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: CDS

LOCATION: (361)..(1605)

US-10-293-244-576

Query Match

Best Local Similarity 0.7%; Score 19; DB 6; Length 1646;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCGGG 75
|||||
DB 22 GTGACCCACGCGTCGGG 40

```
RESULT 68
US-10-273-573-4643/C
; Sequence 4643, Application US/10273573
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: US/10/273,573
; CURRENT FILING DATE: 2002-10-18
; PRIOR APPLICATION NUMBER: 09/522,929
; PRIOR FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: 09/770,160
; PRIOR FILING DATE: 2001-01-26
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 4643
; LENGTH: 1649
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (737)..(1240)
; OTHER INFORMATION: 28% homologous to Drosophila melanogaster hairless
; OTHER INFORMATION: protein,accession number M95192,Smith-Waterman Score=84.
US-10-273-573-4643

Query Match
Best Local Similarity 100.0%; Score 19; DB 6; Length 1649;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 GTGTGACCCACGCGTCCG 73
Db 1630 GTGTGACCCACGCGTCCG 1612

RESULT 69
US-10-357-930-24297
; Sequence 24297, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24297
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-24297

Query Match
Best Local Similarity 100.0%; Score 19; DB 6; Length 1719;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGG 75
Db 2 GTGACCCACGCGTCCGG 20

RESULT 70
US-10-357-930-25866
; Sequence 25866, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25866
; LENGTH: 1719
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-25866

Query Match
Best Local Similarity 100.0%; Score 19; DB 6; Length 1719;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGG 75
Db 2 GTGACCCACGCGTCCGG 20

RESULT 71
US-10-293-244-297
; Sequence 297, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: NOVEL Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
```

```

; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO: 297
; LENGTH: 1783
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (78) ..(1214)
US-10-293-244-297

Query Match          0.7%; Score 19; DB 6; Length 1783;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTGACCCACGCGTCCGGG 75
Db      43 GTGACCCACGCGTCCGGG 61

RESULT 72
US-10-357-930-25092/C
; Sequence 25092, Application US/10357930
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 25092
; LENGTH: 1795
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-25092

Query Match          0.7%; Score 19; DB 6; Length 1795;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTGACCCACGCGTCCGGG 75
Db      1790 GTGACCCACGCGTCCGGG 1772

RESULT 73
US-10-357-930-23084
; Sequence 23084, Application US/10357930
```

```

; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 23084
; LENGTH: 1890
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-23084

Query Match          0.7%; Score 19; DB 6; Length 1890;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      57 GTGACCCACGCGTCCGGG 75
Db      9 GTGACCCACGCGTCCGGG 27

RESULT 74
US-10-293-244-751
; Sequence 751, Application US/10293244
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y. Tom et al
; TITLE OF INVENTION: Novel Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-029
; CURRENT APPLICATION NUMBER: US/10/293,244
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: Not Yet Assigned
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 09/728,422
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: 09/693,325
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 09/663,561
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 09/654,936
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 09/620,325
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/598,075
; PRIOR FILING DATE: 2000-06-20
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 3960
; SOFTWARE: Custom
; SEQ ID NO: 751
; LENGTH: 1930
```

```
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (238)..(1512)
US-10-293-244-751
```

```
Query Match      0.7%; Score 19; DB 6; Length 1930;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      57 GTGACCCACGCGTCCGGG 75
      |||||||
Db      28 GTGACCCACGCGTCCGGG 46
```

RESULT 75

```
US-10-357-930-23024
Sequence 23024, Application US/10357930
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Endege, Wilson
APPLICANT: Monahan, John
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
FILE OF INVENTION: HUMAN PROSTATE CANCER
FILE REFERENCE: MRI-007BCN
CURRENT APPLICATION NUMBER: US/10/357,930
PRIOR FILING DATE: 2003-02-04
PRIOR APPLICATION NUMBER: 09/785,276
PRIOR FILING DATE: 2003-02-16
PRIOR APPLICATION NUMBER: 60/183,319
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 60/189,862
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: 60/207,454
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/211,314
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/219,007
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/255,281
PRIOR FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 62232
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 23024
LENGTH: 1986
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: 1, 1986
OTHER INFORMATION: n = A,T,C or G
US-10-357-930-23024
```

```
Query Match      0.7%; Score 19; DB 6; Length 1986;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      57 GTGACCCACGCGTCCGGG 75
      |||||||
Db      10 GTGACCCACGCGTCCGGG 28
```

Search completed: August 1, 2003, 04:25:55
Job time : 8964 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2003, 19:27:01 / Search time 181 seconds
(without alignments)
6606.115 Million cell updates/sec

Title: US-09-774-490-1

Perfect score: 2709
Sequence: 1 aatctttattatcgcgtg.....agcttttttccctatacc 2709

Scoring table: OLIGO NUC
Gapop 60.0, Gapext 60.0

Searched: 569978 seqs, 220691566 residues

Word size: 0

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database:

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfltest.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	2508	92.6	2601 1	US-08-121-713D-53 Sequence 53, Appl
2	2508	92.6	2601 1	US-08-835-268-53 Sequence 53, Appl
3	2508	92.6	2601 2	US-09-060-692-53 Sequence 53, Appl
4	2508	92.6	2601 3	US-08-833-391-53 Sequence 53, Appl
5	2508	92.6	2601 4	US-09-060-610-53 Sequence 53, Appl
6	2508	92.6	2601 5	PCT-US94-10151A-53 Sequence 53, Appl
7	1253	46.3	1481 1	US-08-136-922-1 Sequence 1, Appl
8	31	1.1	7160 4	US-08-786-531B-5 Sequence 5, Appl
9	31	1.1	7235 4	US-08-786-531B-6 Sequence 5, Appl
10	22	0.8	7352 4	US-08-786-531B-4 Sequence 4, Appl
11	22	0.8	7353 4	US-08-786-531B-1 Sequence 1, Appl
12	21	0.8	1062 4	US-09-391-741A-23 Sequence 23, Appl
13	21	0.8	1062 4	US-09-391-741A-33 Sequence 23, Appl
14	21	0.8	1181 4	US-09-620-312D-33 Sequence 33, Appl
15	21	0.8	1181 4	US-08-984-288-1 Sequence 1, Appl
16	21	0.8	1572 4	US-09-620-312D-664 Sequence 664, App
17	21	0.8	1661 2	US-08-815-176-2 Sequence 2, Appl
18	21	0.8	1661 4	US-09-197-344-2 Sequence 2, Appl
19	21	0.8	2026 4	US-09-324-455-1 Sequence 1, Appl
20	20	0.7	30 3	US-08-850-961-31 Sequence 31, Appl
21	20	0.7	30 3	US-09-132-541-5 Sequence 5, Appl
22	20	0.7	30 4	US-09-479-776-31 Sequence 31, Appl
23	20	0.7	840 4	US-09-244-111-5 Sequence 5, Appl
24	20	0.7	1059 4	US-09-391-741A-9 Sequence 9, Appl
25	20	0.7	1168 4	US-09-620-312D-562 Sequence 562, App
26	20	0.7	1279 3	US-08-985-950-5 Sequence 5, Appl
27	20	0.7	1279 4	US-09-546-049-5 Sequence 5, Appl

28	20	0.7	1338 4	US-09-336-536-1 Sequence 1, Appl
29	20	0.7	1503 3	US-08-999-774A-11 Sequence 11, Appl
30	20	0.7	1772 2	US-09-482-273-96 Sequence 96, Appl
31	20	0.7	1920 4	US-09-620-312D-7 Sequence 7, Appl
32	20	0.7	2214 4	US-09-489-847-113 Sequence 113, App
33	20	0.7	2225 4	US-09-620-312D-1046 Sequence 1046, App
34	20	0.7	2227 4	US-09-489-847-30 Sequence 30, Appl
35	20	0.7	4137 3	US-09-221-235-1 Sequence 1, Appl
36	20	0.7	4137 3	US-09-221-928-1 Sequence 1, Appl
37	20	0.7	4137 3	US-09-221-577-1 Sequence 1, Appl
38	20	0.7	4137 3	US-09-221-236-1 Sequence 1, Appl
39	20	0.7	4137 3	US-09-221-416-1 Sequence 1, Appl
40	20	0.7	4137 3	US-09-221-245-1 Sequence 1, Appl
41	20	0.7	4137 3	US-09-163-115-1 Sequence 1, Appl
42	20	0.7	4137 3	US-09-221-528-1 Sequence 1, Appl
43	20	0.7	4137 3	US-09-593-553-1 Sequence 1, Appl
44	20	0.7	4137 3	US-09-221-237-1 Sequence 1, Appl
45	20	0.7	5865 3	US-08-654-737B-3 Sequence 3, Appl
46	20	0.7	6700 4	US-09-654-449-1 Sequence 1, Appl
47	20	0.7	8518 4	US-09-654-449-3 Sequence 3, Appl
48	20	0.7	8540 4	US-08-487-283A-4 Sequence 4, Appl
49	20	0.7	8540 5	PCT-US96-05611A-12 Sequence 12, Appl
50	19	0.7	256 4	US-09-313-294A-2526 Sequence 2526, App
51	19	0.7	342 3	US-09-461-697-136 Sequence 136, App
52	19	0.7	571 4	US-09-620-312D-920 Sequence 920, App
53	19	0.7	581 4	US-09-512-363-9 Sequence 9, Appl
54	19	0.7	581 4	US-09-176-200-9 Sequence 9, Appl
55	19	0.7	648 2	US-08-801-972-2 Sequence 2, Appl
56	19	0.7	648 3	US-09-178-881-2 Sequence 2, Appl
57	19	0.7	780 4	US-09-205-258-63 Sequence 63, Appl
58	19	0.7	831 4	US-09-904-615-25 Sequence 25, Appl
59	19	0.7	852 3	US-09-461-697-1 Sequence 1, Appl
60	19	0.7	9037 4	US-09-257-179-20 Sequence 20, Appl
61	19	0.7	1007 4	US-09-512-363-3 Sequence 3, Appl
62	19	0.7	1007 4	US-09-176-200-3 Sequence 3, Appl
63	19	0.7	1024 4	US-09-620-312D-866 Sequence 866, App
64	19	0.7	1037 4	US-09-489-847-112 Sequence 112, App
65	19	0.7	1116 4	US-09-372-422A-41 Sequence 41, Appl
66	19	0.7	1221 4	US-09-620-312D-162 Sequence 162, App
67	19	0.7	1281 4	US-09-620-312D-972 Sequence 972, App
68	19	0.7	1332 4	US-09-620-312D-196 Sequence 196, App
69	19	0.7	1354 4	US-09-620-312D-308 Sequence 308, App
70	19	0.7	1368 4	US-09-620-312D-871 Sequence 871, App
71	19	0.7	1390 4	US-09-205-258-124 Sequence 124, App
72	19	0.7	1411 3	US-08-964-127-5 Sequence 5, Appl
73	19	0.7	1411 4	US-09-466-692-5 Sequence 5, Appl
74	19	0.7	1411 4	US-10-000-273-5 Sequence 5, Appl
75	19	0.7	1428 3	US-09-118-442-5 Sequence 5, Appl
76	19	0.7	1428 3	US-09-677-064-5 Sequence 5, Appl
77	19	0.7	1448 4	US-09-130-491-9 Sequence 9, Appl
78	19	0.7	1466 4	US-09-620-312D-914 Sequence 914, App
79	19	0.7	1503 4	US-09-620-312D-588 Sequence 588, App
80	19	0.7	1509 4	US-09-620-312D-722 Sequence 722, App
81	19	0.7	1695 3	US-09-089-879-1 Sequence 1, Appl
82	19	0.7	1695 3	US-09-488-260-1 Sequence 1, Appl
83	19	0.7	1700 3	US-08-988-856B-3 Sequence 3, Appl
84	19	0.7	1700 5	PCT-US96-01806-3 Sequence 3, Appl
85	19	0.7	1939 4	US-09-719-083A-3 Sequence 3, Appl
86	19	0.7	1989 4	US-09-620-312D-837 Sequence 837, App
87	19	0.7	2045 4	US-09-620-312D-589 Sequence 589, App
88	19	0.7	2062 4	US-09-620-312D-170 Sequence 170, App
89	19	0.7	2095 4	US-09-620-312D-748 Sequence 748, App
90	19	0.7	2102 4	US-09-620-312D-487 Sequence 487, App
91	19	0.7	2210 4	US-09-620-312D-1009 Sequence 1009, App
92	19	0.7	2243 4	US-09-620-312D-1008 Sequence 1008, App
93	19	0.7	2254 4	US-09-620-312D-73 Sequence 73, Appl
94	19	0.7	2351 4	US-09-620-312D-1007 Sequence 1007, App
95	19	0.7	2504 4	US-09-620-312D-778 Sequence 778, App
96	19	0.7	2525 3	US-09-342-648-1 Sequence 1, Appl
97	19	0.7	2692 4	US-09-996-243-298 Sequence 298, App
98	19	0.7	2702 3	US-09-461-697-359 Sequence 359, App
99	19	0.7	2874 4	US-09-620-312D-10 Sequence 10, Appl
100	19	0.7	3045 4	US-09-620-312D-9 Sequence 9, Appl

ALIGNMENTS

```
RESULT 1
US-08-121-713D-53
; Sequence 53, Application US/08121713D
; Patent No. 5639856
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Matthes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713D
; FILING DATE: 13-SEP-1993
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Oseman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415)343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2601 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..2331
; US-08-121-713D-53

Query Match      92.6%; Score 2508; DB 1; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      193  CTGAGCATGGGCTGTTAACTAGAGTGTCTGCTTTCTGGGGAGTATTACTTAACGC 252
DB      9   CTGAGCATGGGCTGTTAACTAGAGTGTCTGCTTTCTGGGGAGTATTACTTAACGC 68
QY      253  AAGAGCAACTATCAGAAATGGAGAGAACATGTGCCAAGCTGAATTAATCTTAACAAGA 312
DB      69  AAGAGCAACTATCAGAAATGGAGAGAACATGTGCCAAGCTGAATTAATCTTAACAAGA 128
QY      313  AATGTTGGAATCAACAATGTATGATCACTTTCAATGGCTGGCAACAGCTCCAGTTATCA 372
DB      129  AATGTTGGAATCAACAATGTATGATCACTTTCAATGGCTGGCAACAGCTCCAGTTATCA 188
QY      373  TACCTTCCTTTTGATGAGAGAGAGTGTGCTGTATGTTGAGCAAAAGGATCACATATT 432
DB      189  TACCTTCCTTTTGATGAGAGAGAGTGTGCTGTATGTTGAGCAAAAGGATCACATATT 248
```

```
QY      433  TTCAATGCACTGGTTAAATCAAGATTTTCAAAAGATGTGTGGCCAGTATCTTAAC 492
DB      249  TTCAATGCACTGGTTAAATCAAGATTTTCAAAAGATGTGTGGCCAGTATCTTAAC 308
QY      493  CAGAAAGATGATTCAGATGGGCTGGAAAAGACATCTGTGAAAAGATGTCTAATTTAT 552
DB      309  CAGAAAGATGATTCAGATGGGCTGGAAAAGACATCTGTGAAAAGATGTCTAATTTAT 368
QY      553  CAAAGTACTTAAGGATTAATCAAGACTCACTTGTAGCCCTGTGGACGGGGGCTTTTCA 612
DB      369  CAAAGTACTTAAGGATTAATCAAGACTCACTTGTAGCCCTGTGGACGGGGGCTTTTCA 428
QY      613  TCCAAATTTGCACTACATTAAGATGACATCATCTTGAGAGCAATATTTTAAGCTGGA 672
DB      429  TCCAAATTTGCACTACATTAAGATGACATCATCTTGAGAGCAATATTTTAAGCTGGA 488
QY      673  GAATCAATTTTGAAGAAACGGCCGTGGGAAAGATTCATATGACCTTAAGCTGTACAGC 732
DB      489  GAATCAATTTTGAAGAAACGGCCGTGGGAAAGATTCATATGACCTTAAGCTGTACAGC 548
QY      733  ATCCCTTTAATATGATGAGAAATTAATTAATCTGGAACCTGACGTGATTTATGGGGCGAGA 792
DB      549  ATCCCTTTAATATGATGAGAAATTAATTAATCTGGAACCTGACGTGATTTATGGGGCGAGA 608
QY      793  CTTGCTATCTTCCGAACCTTTGGGACACACACCCCAATCAGACAGACATGATTC 852
DB      609  CTTGCTATCTTCCGAACCTTTGGGACACACACCCCAATCAGACAGACATGATTC 668
QY      853  CAGGTGGCTCAATGATCCAAAGTTTATGAGCCCACTCATCTCAGAGATGACAAATCC 912
DB      669  CAGGTGGCTCAATGATCCAAAGTTTATGAGCCCACTCATCTCAGAGATGACAAATCC 728
QY      913  TGAAGATGACAAAGATATCTTTTCTTCCGGAAGATTCATATGATGAGAAACATCTGG 972
DB      729  TGAAGATGACAAAGATATCTTTTCTTCCGGAAGATTCATATGATGAGAAACATCTGG 788
QY      973  AAAAGCTACACGCTAGATAGGTCAATATGCAAGATGACCTTTGAGGGGCAACAGAG 1032
DB      789  AAAAGCTACACGCTAGATAGGTCAATATGCAAGATGACCTTTGAGGGGCAACAGAG 848
QY      1033  TCTGTGAATTAATGAGCAACATTTCTCAAGCTGTCTGATTTGCTCAGAGCCAGATTC 1092
DB      849  TCTGTGAATTAATGAGCAACATTTCTCAAGCTGTCTGATTTGCTCAGAGCCAGATTC 908
QY      1093  AAATGGCACTGACACTCTTTTGAATGAACCTCAGAGATGATTTCTTAATGAATTTAAGA 1152
DB      909  AAATGGCACTGACACTCTTTTGAATGAACCTCAGAGATGATTTCTTAATGAATTTAAGA 968
QY      1153  TCTTAAAAATCAGTTGTATATGAGATGTTTACGACTTCAGTAAACATTTTCAAGGATC 1212
DB      969  TCTTAAAAATCAGTTGTATATGAGATGTTTACGACTTCAGTAAACATTTTCAAGGATC 1028
QY      1213  AGCCGTGTATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1272
DB      1029  AGCCGTGTATATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1088
QY      1273  CAGGATGAGACCAACTATCAATGGGTGCTTATCAAGAAAGATGCCCTATCAACGGCC 1332
DB      1089  CAGGATGAGACCAACTATCAATGGGTGCTTATCAAGAAAGATGCCCTATCAACGGCC 1148
QY      1333  AGGAACTTGTCCAGCAAAACATTTGTGTGTGTTGACTTCAAAAGACCTTCTGATGA 1392
DB      1149  AGGAACTTGTCCAGCAAAACATTTGTGTGTGTTGACTTCAAAAGACCTTCTGATGA 1208
QY      1393  TGTATTAACCTTTTGCAAGAGTATCTCAGGCAATGATCAATCAATGTTTCTTAAGACAA 1452
DB      1209  TGTATTAACCTTTTGCAAGAGTATCTCAGGCAATGATCAATCAATGTTTCTTAAGACAA 1268
QY      1453  TCGCCCAATATGATCAAAAACGATGTAATTAATTAATTAATTAATTAATTAATTAATTA 1512
DB      1269  TCGCCCAATATGATCAAAAACGATGTAATTAATTAATTAATTAATTAATTAATTAATTA 1328
QY      1513  AGTGAATGACAGAAATGACAGTATGATGTTATGTTATGAGAAACAGATGTTGGACCGT 1572
```



```

Db 1329 AGTGAATCAGAGAGATGACAGTATGATATGTTATCGGAAAGATGTTGGACCGT 1388
Qy 1573 TCTTAAAGTAGTTCAATTCCTAAGAGAGACTTGGATATATTAGAGAGGTTCTGCTGA 1632
Db 1389 TCTTAAAGTAGTTCAATTCCTAAGAGAGACTTGGATATATTAGAGAGGTTCTGCTGA 1448
Qy 1633 AGAATGACAGTTCCTCGGAAACCGACTGCTATTTCAAGCAATGAGCTTTCACATAAGCA 1692
Db 1449 AGAATGACAGTTCCTCGGAAACCGACTGCTATTTCAAGCAATGAGCTTTCACATAAGCA 1508
Qy 1693 GCAACACTATATATTTGTTTCAACGGCTGGGTTGCCAAGCTCCCTTACACCGGTGTA 1752
Db 1509 GCAACACTATATATTTGTTTCAACGGCTGGGTTGCCAAGCTCCCTTACACCGGTGTA 1568
Qy 1753 TATTTAGGGGAAAGGCTGCTGAGTGTGCTGCGCCGAGACCCCTTACTGCTGGGA 1812
Db 1569 TATTTAGGGGAAAGGCTGCTGAGTGTGCTGCGCCGAGACCCCTTACTGCTGGGA 1628
Qy 1813 TGGTTCGTCAGTTCCTCGCTATTTTCCACTGCAAGAGACGCAAGACGACAGATAT 1872
Db 1629 TGGTTCGTCAGTTCCTCGCTATTTTCCACTGCAAGAGACGCAAGACGACAGATAT 1688
Qy 1873 AAGAAATGAGACCCACTGACTCACTGTTCAAGCTTACACATGATTAACCATGSCCA 1932
Db 1689 AAGAAATGAGACCCACTGACTCACTGTTCAAGCTTACACATGATTAACCATGSCCA 1748
Qy 1933 CAGCCCTGAAGAGAAATCATCTATGCTAGTAGAGAAATAGTAGACATTTTGGATGAG 1992
Db 1749 CAGCCCTGAAGAGAAATCATCTATGCTAGTAGAGAAATAGTAGACATTTTGGATGAG 1808
Qy 1993 TCCGAAGTCGAGAGAGCGCTGCTATTTGCAATTCAGAGGCGCAATGGAAGCGAA 2052
Db 1809 TCCGAAGTCGAGAGAGCGCTGCTATTTGCAATTCAGAGGCGCAATGGAAGCGAA 1868
Qy 2053 AGAAGAGATCAAGTGTATGATCATATCATGACAGACATCAAGGCTTCTGCTACGTAG 2112
Db 1869 AGAAGAGATCAAGTGTATGATCATATCATGACAGACATCAAGGCTTCTGCTACGTAG 1928
Qy 2113 TCTTCAACAGAAAGATTGAGCAATTAACCTGCGCATGCGGTGGAACATGGGCTCATCA 2172
Db 1929 TCTTCAACAGAAAGATTGAGCAATTAACCTGCGCATGCGGTGGAACATGGGCTCATCA 1988
Qy 2173 AACTCTTCTTAAAGTAACCTCGAAGTATTGACACAGCAATTTGGAAGAACTTCTTCA 2232
Db 1989 AACTCTTCTTAAAGTAACCTCGAAGTATTGACACAGCAATTTGGAAGAACTTCTTCA 2048
Qy 2233 TAAAGATGATGAGATGAGCTCTTACAGCAAGAAATGTCATATAGCATGACACTTAC 2292
Db 2049 TAAAGATGATGAGATGAGCTCTTACAGCAAGAAATGTCATATAGCATGACACTTAC 2108
Qy 2293 CCAAGAGGTCTGCTACAGAGACTTCAATGAGCTCATCAACCCCAATCTCAACAGAT 2352
Db 2109 CCAAGAGGTCTGCTACAGAGACTTCAATGAGCTCATCAACCCCAATCTCAACAGAT 2168
Qy 2353 GGAATGAGTCTGCTGAAACAAGTTTGGAAAGGAGCCGAAACAGCTCGGCAAGGCCAG 2412
Db 2169 GGAATGAGTCTGCTGAAACAAGTTTGGAAAGGAGCCGAAACAGCTCGGCAAGGCCAG 2228
Qy 2413 ACATATCCCAAGGAAACAGTAACAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAA 2472
Db 2229 ACATATCCCAAGGAAACAGTAACAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAA 2288
Qy 2473 CAGAGAGCCCAAGAAATTTGAGAGGGCACCAAGAGTGTCTGAGTGAATTAACCTTGA 2532
Db 2289 CAGAGAGCCCAAGAAATTTGAGAGGGCACCAAGAGTGTCTGAGTGAATTAACCTTGA 2348
Qy 2533 AACCTCAAAAGTAGAAGTGTCTAGACATAATCTGGAAGAAACAAATGCAATATACAT 2592
Db 2349 AACCTCAAAAGTAGAAGTGTCTAGACATAATCTGGAAGAAACAAATGCAATATACAT 2408
Qy 2593 GAACCTTTTTCATGCGATTAATGTGATGTTTCAATGAGTGGAAATTCAGCTGAGTTTCA 2652

```

```

Db 2409 GAACCTTTTTCATGCGATTAATGTGATGTTTACATGAGTGGAAATTCAGCTGAGTTTCA 2468
Qy 2653 CCAATTAATAATTAATCCATGAGTAACCTTCTTAATAGGCTTTT 2700
Db 2469 CCAATTAATAATTAATCCATGAGTAACCTTCTTAATAGGCTTTT 2516

RESULT 2
US-08-835-268-53
; Sequence 53, Application US/08835268
; Patent No. 5807826
; GENERAL INFORMATION:
; APPLICANT: Goodman, Corey S.
; APPLICANT: Kolodkin, Alex L.
; APPLICANT: Mathes, David
; APPLICANT: Bentley, David R.
; APPLICANT: O'Connor, Timothy
; TITLE OF INVENTION: The Semaphorin Gene Family
; NUMBER OF SEQUENCES: 100
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
; STREET: 268 Bush Street, Suite 3200
; CITY: San Francisco
; STATE: CA
; COUNTRY: USA
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/835,268
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,713
; FILING DATE: 13-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Osman, Richard A.
; REGISTRATION NUMBER: 36,627
; REFERENCE/DOCKET NUMBER: B94-002-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 343-4341
; TELEFAX: (415) 343-4342
; TELEX:
; INFORMATION FOR SEQ ID NO: 53:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2601 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 16..2331
; US-08-835-268-53

Query Match 92.6%; Score 2508; DB 1; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY	373	TACCTTCCTTTTGGATGAGGAACGGAGTAGGCGTGTATGTTGGACGAAGAATCACTATT	432
Db	189	TACCTTCCTTTTGGATGAGGAACGGAGTAGGCGTGTATGTTGGACGAAGAATCACTATT	248
QY	433	TTCAATTCGACCTGGTGTATATCAAGGATTTTCAAAAGATTGTGTGGCCAGTATCTTACAC	492
Db	249	TTCAATTCGACCTGGTGTATATTCAGAGATTTTCAAAAGATTGTGTGGCCAGTATCTTACAC	308
QY	493	CAGAAGAGATGAATGCAAGTGGGCTGGAAGAAGACATCTGGAAGAATGTGCTAATTTTCAT	552
Db	309	CAGAAGAGATGAATGCAAGTGGGCTGGAAGAAGACATCTGGAAGAATGTGCTAATTTTCAT	368
QY	553	CAGAGACTTAAGGGCAATATTAACAGCTCACTGTAGCGCTGTGGAAACGGGGCTTTTCA	612
Db	369	CAGAGACTTAAGGGCAATATTAACAGCTCACTGTAGCGCTGTGGAAACGGGGCTTTTCA	428
QY	613	TCCAAATTTGACCTACATTTGAATTTGGAACATCATCTGAGAGCAATATTTTAAAGCTGGA	672
Db	429	TCCAAATTTGACCTACATTTGAATTTGGAACATCATCTGAGAGCAATATTTTAAAGCTGGA	488
QY	673	GAACTCACATTTTGAAGAACGGCCGTGGGAAGAGTCCATATGACCTTAAGCTGTGCAACAC	732
Db	489	GAACTCACATTTTGAAGAACGGCCGTGGGAAGAGTCCATATGACCTTAAGCTGTGCAACAC	548
QY	733	ATCCCTTTTAATATGATGAGGAATTAATCTCTGGAATCTGACAGCTGATTTTATGGGGGAGA	792
Db	549	ATCCCTTTTAATATGATGAGGAATTAATCTCTGGAATCTGACAGCTGATTTTATGGGGGAGA	608
QY	793	CTTTGCTATCTTCCGAACCTTTGGGACCAACCAACCAATCAGGACAGACAGCATATTC	852
Db	609	CTTTGCTATCTTCCGAACCTTTGGGACCAACCAACCAATCAGGACAGACAGCATATTC	668
QY	853	CAGGTGGCTCAATGATCCAAAGTTCAATAGTCCCACTTCATCTCAGAAGTGAACAATCC	912
Db	669	CAGGTGGCTCAATGATCCAAAGTTCAATAGTCCCACTTCATCTCAGAAGTGAACAATCC	728
QY	913	TGAAGATGCAAGATATACCTTTTCTCGGTGAAATGCAATATGAGGGAACACTCTGG	972
Db	729	TGAAGATGCAAGATATACCTTTTCTCGGTGAAATGCAATATGAGGGAACACTCTGG	788
QY	973	AAAAGCTACTCAGCTAGATAGTGCAGATATGCAAGATGACTTTGAGGGCACAGAAG	1033
Db	789	AAAAGCTACTCAGCTAGATAGTGCAGATATGCAAGATGACTTTGAGGGCACAGAAG	848
QY	1033	TCTGTGAATTAATGAGCAACAATCTCTCAAAGCTGCTGTGATTTGCTCAGTGCCAGTCC	1093
Db	849	TCTGTGAATTAATGAGCAACAATCTCTCAAAGCTGCTGTGATTTGCTCAGTGCCAGTCC	908
QY	1093	AAATGGCATTTGACACTCTTTTGAATGACATGACGAGATGATTTCTTAATGAACTTTAAGA	1153
Db	909	AAATGGCATTTGACACTCTTTTGAATGACATGACGAGATGATTTCTTAATGAACTTTAAGA	968
QY	1153	TCTTAATAATCCAGTTGTATATGAGAGTGTTAACATTCACAGTAACATTTCAAGGATC	1213
Db	969	TCTTAATAATCCAGTTGTATATGAGAGTGTTAACATTCACAGTAACATTTCAAGGATC	1028
QY	1213	AGCCGTGTATATGATAGCATGATGATGAGAGAGGTGTTCTTGTGTCATATGACCA	1273
Db	1029	AGCCGTGTATATGATAGCATGATGATGAGAGAGGTGTTCTTGTGTCATATGACCA	1088
QY	1273	CAGGATGGAACCACTATCAATGGGTGCTTAACGAAGAAGAGTCCCTATTCACAGGCC	1333
Db	1089	CAGGATGGAACCACTATCAATGGGTGCTTAACGAAGAAGAGTCCCTATTCACAGGCC	1144
QY	1333	AGGAACCTGTCCAGCAAAACATTTGTGTGTTTGACTCTAACAAGGACCTTCGTATGA	1393
Db	1149	AGGAACCTGTCCAGCAAAACATTTGTGTGTTTGACTCTAACAAGGACCTTCGTATGA	1208
QY	1393	TGTTATTAACCTTTGCAAGAGTCAATCCAGCAATGTAACAATCAAGTGTTCCTATGAAACA	1453
Db	1209	TGTTATTAACCTTTGCAAGAGTCAATCCAGCAATGTAACAATCAAGTGTTCCTATGAAACA	1268

QY	1453	TGCGCCCATAGTATGCAAAAGGATGTAAATTTATCAATTTACAAAATTTGCTGTAGACCG	1512
Db	1269	TGCGCCCAATAGGATCAAAACGGATGTAAATTAATCAATTTACAAAATTTGCTGTAGACCG	1328
QY	1513	AGTGGATGCAGAAAGATGGACAGTATGATGTTATGTTATTCGGAAACAGATGTTGGGACCGT	1572
Db	1329	AGTGGATGCAGAAAGATGGACAGTATGATGTTATGTTATTCGGAAACAGATGTTGGGACCGT	1388
QY	1573	TCTTAAAGTAGTTCAATTCCTTAAGAGACCTTGGATGATGATTTAGAAAGGTTCTGTCTGGA	1632
Db	1389	TCTTAAAGTAGTTCAATTCCTTAAGAGACCTTGGATGATGATTTAGAAAGGTTCTGTCTGGA	1448
QY	1633	AGAAATGACAGTTTTTCGGGAAACCGACCTGCTATTTTCAGACATTGGACCTTCCACTAAGCA	1692
Db	1449	AGAAATGACAGTTTTTCGGGAAACCGACCTGCTATTTTCAGACATTGGACCTTCCACTAAGCA	1508
QY	1693	GCAACAACTATATATTTGGTTCAACGGCTGGGGGTTGCCAGCTCCCTTTACACGGGTGGA	1752
Db	1509	GCAACAACTATATATTTGGTTCAACGGCTGGGGGTTGCCAGCTCCCTTTACACGGGTGGA	1568
QY	1753	TATTTACCGGAAAGCGTGTGCTGATGTTGTGCTGTGCGCCGAGACCCCTTACTGTGCTTGGGA	1812
Db	1569	TATTTACCGGAAAGCGTGTGCTGATGTTGTGCTGTGCGCCGAGACCCCTTACTGTGCTTGGGA	1628
QY	1813	TGTTTCTGCATGTTCTGTGCTATATTTTCCCACTGTCAAGACGCAACAGACACAATAT	1872
Db	1629	TGTTTCTGCATGTTCTGTGCTATATTTTCCCACTGTCAAGACGCAACAGACACAATAT	1688
QY	1873	AAGAAATGGAGAACCCACTGACTCACTGTTCAGACTTACACCATATATATCAACATGGCCA	1932
Db	1689	AAGAAATGGAGAACCCACTGACTCACTGTTCAGACTTACACCATATATATCAACATGGCCA	1748
QY	1933	CAGCCCTGAAGAGAAATCATCTATGTTGAGAAATAGTACACATTTTGGAAATGCAG	1992
Db	1749	CAGCCCTGAAGAGAAATCATCTATGTTGAGAAATAGTACACATTTTGGAAATGCAG	1808
QY	1993	TCCGAAGTGCAGAGAGCGCTGTGCTATATTTGGCAATTCAGAGCGGAAATGAAGCGGAA	2052
Db	1809	TCCGAAGTGCAGAGAGCGCTGTGCTATATTTGGCAATTCAGAGCGGAAATGAAGCGGAA	1868
QY	2053	AGAAGAGTCAAGAGTGAATCATATCATATGATGAGACAGATCAAGAGCCCTTCGTATCGTAA	2112
Db	1869	AGAAGAGTCAAGAGTGAATCATATCATATGATGAGACAGATCAAGAGCCCTTCGTATCGTAA	1928
QY	2113	TCTACAAAGAGGATTCAGGCAATTAACCTTGTCCATGCGGTGGAGATGGGTTTCATACA	2172
Db	1929	TCTACAAAGAGGATTCAGGCAATTAACCTTGTCCATGCGGTGGAGATGGGTTTCATACA	1988
QY	2173	AACTCTTTTAAGGTAACCTCTGAAAGTCAATTGACACAGAGCATTTGGAAAGACTTTCTTCA	2232
Db	1989	AACTCTTTTAAGGTAACCTCTGAAAGTCAATTGACACAGAGCATTTGGAAAGACTTTCTTCA	2048
QY	2233	TAAAGATGATATGAGATGAGTGTCTTAAGACCAAGAAATGTCCAAATGATGACACTAG	2292
Db	2049	TAAAGATGATATGAGATGAGTGTCTTAAGACCAAGAAATGTCCAAATGATGACACTAG	2108
QY	2293	CCAGAGGCTGTGTATACAGAGACTTCATGACAGCTCATCAACCAACCCCAATCTCAACAGAT	2352
Db	2109	CCAGAGGCTGTGTATACAGAGACTTCATGACAGCTCATCAACCAACCCCAATCTCAACAGAT	2168
QY	2353	GGATGAGTCTGTGTAACAAGTTTGGAAAAGGGAACGAAAACAACGTCGGCAAGGCGCAGG	2412
Db	2169	GGATGAGTCTGTGTAACAAGTTTGGAAAAGGGAACGAAAACAACGTCGGCAAGGCGCAGG	2228
QY	2413	ACATATCCCAAGGAAACGTAAACAATGGAAGCACTTACAGAAATATGAAGAAAGGTATGAAA	2472
Db	2229	ACATATCCCAAGGAAACGTAAACAATGGAAGCACTTACAGAAATATGAAGAAAGGTATGAAA	2288
QY	2473	CAGAGAGCCCAAGAAATTTGAGAGGGACCCAGAGAGTGTCTGAGCTGACTATTAACCTCTAGA	2532
Db	2289	CAGAGAGCCCAAGAAATTTGAGAGGGACCCAGAGAGTGTCTGAGCTGACTATTAACCTCTAGA	2348
QY	2533	AACCTCAAAACAGTAGAAACTTGTCTGACAAATTAACGTGAAGAAAACAATGTCAATATATACAT	2592

Db 9 CTGACGACATGGGCTGGTAACTAGATGTCTGTCTTTCTGGGAGATATCTACAGC 68
Qy 253 AAGACAAACATCATGAAATGGAGAAACAATGGCCAGGCTGAAATATCTACAAAGA 312
Db 69 AAGACAAACATCATGAAATGGAGAAACAATGGCCAGGCTGAAATATCTACAAAGA 128
Qy 313 AATGTTGAATCCAAATGTGATCACTTCAATGGCTTGGCCAAAGCTCCAGTTATCA 372
Db 129 AATGTTGAATCCAAATGTGATCACTTCAATGGCTTGGCCAAAGCTCCAGTTATCA 188
Qy 373 TACCTTCTTTTGGATGAGGAAACGAGTGGCTGATGTTGGAGCAAAAGATCAATAT 432
Db 189 TACCTTCTTTTGGATGAGGAAACGAGTGGCTGATGTTGGAGCAAAAGATCAATAT 248
Qy 433 TTCAATTCGACCTGATTAATATCAAGGATTTTCAAAAGATGTGTGGCCAGTATCTTAAC 492
Db 249 TTCAATTCGACCTGATTAATATCAAGGATTTTCAAAAGATGTGTGGCCAGTATCTTAAC 308
Qy 493 CAGAAAGATGAAATGCAAGTGGCTGAAAGAAAGATCCGAAAGATGTCTAATTTGAT 552
Db 309 CAGAAAGATGAAATGCAAGTGGCTGAAAGAAAGATCCGAAAGATGTCTAATTTGAT 368
Qy 553 CAGGATCTTAAAGGCAATATATCAGACTGCTTGTACCGCTGTGGAACGGGGCTTTTCA 612
Db 369 CAGGATCTTAAAGGCAATATATCAGACTGCTTGTACCGCTGTGGAACGGGGCTTTTCA 428
Qy 613 TCCAAATTTGACCTTCAATTTGAATTTGGAATCTGATCATCTGAGGACATATTTTAACCTGA 672
Db 429 TCCAAATTTGACCTTCAATTTGAATTTGGAATCTGATCATCTGAGGACATATTTTAACCTGA 488
Qy 673 GAACCTCACATTTTGAAGAAACGGCCGTGGAGAGAGTCCATATGACCTTAAGCTGCTGACAGC 732
Db 489 GAACCTCACATTTTGAAGAAACGGCCGTGGAGAGAGTCCATATGACCTTAAGCTGCTGACAGC 548
Qy 733 ATCCCTTTTAATATGATGAGAAATTAATCTCTGGAATCTGACGCTGATTTTAATGGGCGAGA 792
Db 549 ATCCCTTTTAATATGATGAGAAATTAATCTCTGGAATCTGACGCTGATTTTAATGGGCGAGA 608
Qy 793 CTTTCTCATCTTCCGAACTCTTGGGACACCAACCCATTCAGGACAGAGCATGATTC 852
Db 609 CTTTCTCATCTTCCGAACTCTTGGGACACCAACCCATTCAGGACAGAGCATGATTC 668
Qy 853 CAGGTGGCTCAATGATCCAAAGTTCATTAAGTCCCACTCATCTCAGAGAGTGAATCC 912
Db 669 CAGGTGGCTCAATGATCCAAAGTTCATTAAGTCCCACTCATCTCAGAGAGTGAATCC 728
Qy 913 TGAAGATGACAAAGTAACTTTTCTTCGCTGAAATGCAATATGATGAGAAACACTCTGG 972
Db 729 TGAAGATGACAAAGTAACTTTTCTTCGCTGAAATGCAATATGATGAGAAACACTCTGG 788
Qy 973 AAAAGCTACGACGCTAATAATAGTCAGATATGCAATATGCAATTTGAGAGGCAACAGAG 1032
Db 789 AAAAGCTACGACGCTAATAATAGTCAGATATGCAATATGCAATTTGAGAGGCAACAGAG 848
Qy 1033 TCTGTGTAATTAATGAGCAACATCTCTCAAAAGCTGCTGATTTTCTCAGTGGCCAGTCC 1092
Db 849 TCTGTGTAATTAATGAGCAACATCTCTCAAAAGCTGCTGATTTTCTCAGTGGCCAGTCC 908
Qy 1093 AAATGGCAATGACACTATTTTGTATGAACTGCAAGATGTATTTCTTAATGAACTTTAAAGA 1152
Db 909 AAATGGCAATGACACTATTTTGTATGAACTGCAAGATGTATTTCTTAATGAACTTTAAAGA 968
Qy 1153 TCCCTAAATATCAGTGTATATATGAGTGTATTAACGATCCAGTAACTTTTCAAGGATC 1212
Db 969 TCCCTAAATATCAGTGTATATATGAGTGTATTAACGATCCAGTAACTTTTCAAGGATC 1028
Qy 1213 AGCCGTGTATATGATAGCATAGATGATGAGAGGCTGTTCTTGGTCCATATGCCCA 1272
Db 1029 AGCCGTGTATATGATAGCATAGATGATGAGAGGCTGTTCTTGGTCCATATGCCCA 1088
Qy 1273 CAGGATGAGACCAACTATCAATGGGTGCTTATTAAGAAAGAGTCCCTTATCCAGGCTC 1332

Db 1089 CAGGATGAGACCAACTATCAATGAGTCCCTTATCAAGAAAGATCCCTATCCAGGCTC 1148
Qy 1333 AGGAACCTGTCCGACCAAAACATTTGGTGTGTTGACTATCAAAAGACCTTCCGATGA 1392
Db 1149 AGGAACCTGTCCGACCAAAACATTTGGTGTGTTGACTATCAAAAGACCTTCCGATGA 1208
Qy 1393 TGTATTAACCTTTCAAGAAATCATCCAGCCATGTACATCAGTGTTCCTATGAACAA 1452
Db 1209 TGTATTAACCTTTCAAGAAATCATCCAGCCATGTACATCAGTGTTCCTATGAACAA 1268
Qy 1453 TCGCCCAATAGTATCAAAACGAGTGAATTAATTAATTAACAAATTTGTGTAGACCG 1512
Db 1269 TCGCCCAATAGTATCAAAACGAGTGAATTAATTAATTAACAAATTTGTGTAGACCG 1328
Qy 1513 AGTGAATGACAAATGAGACAGTATGATATGTTATATGGAACAGATGTTGGAGCCGT 1572
Db 1329 AGTGAATGACAAATGAGACAGTATGATATGTTATATGGAACAGATGTTGGAGCCGT 1388
Qy 1573 TCTTAAAGTATGTTCAATCTCTTAAGAGACTTGGTATGATTTAGAAAGGTTCTGTGGA 1632
Db 1389 TCTTAAAGTATGTTCAATCTCTTAAGAGACTTGGTATGATTTAGAAAGGTTCTGTGGA 1448
Qy 1633 AGAATGACAGTTTTGGGAAACCGACTGCTATTTACGAATGAGAGCTTTCCTAAGCA 1692
Db 1449 AGAATGACAGTTTTGGGAAACCGACTGCTATTTACGAATGAGAGCTTTCCTAAGCA 1508
Qy 1693 GCACCACTATATATGTTGTTCAAGCGCTGGGCTTGGCCAGTCCCTTACACCGGTGGA 1752
Db 1509 GCACCACTATATATGTTGTTCAAGCGCTGGGCTTGGCCAGTCCCTTACACCGGTGGA 1568
Qy 1753 TATTTAGGGGAAACGTTGTGCTGAGTGTGCTGCGCCGAGAACCTTACTGTCTGGGA 1812
Db 1569 TATTTAGGGGAAACGTTGTGCTGAGTGTGCTGCGCCGAGAACCTTACTGTCTGGGA 1628
Qy 1813 TGTGTTCTGATGTTCTGCTATTTTCCCATCTGCAAGAGACGCAAGAGCAAGATAT 1872
Db 1629 TGTGTTCTGATGTTCTGCTATTTTCCCATCTGCAAGAGACGCAAGAGCAAGATAT 1688
Qy 1873 AAGAAATGAGAGACCCAGCTGACTCATGTTACAGATTAACCATATATATCAACATGGCCA 1932
Db 1689 AAGAAATGAGAGACCCAGCTGACTCATGTTACAGATTAACCATATATATCAACATGGCCA 1748
Qy 1933 CAGCCCTGAGAGAGAAATCATATGATGATGAGAAATAGTATGACATTTTGTGAATCAG 1992
Db 1749 CAGCCCTGAGAGAGAAATCATATGATGATGAGAAATAGTATGACATTTTGTGAATCAG 1808
Qy 1993 TCCGAATGCGAGAGACGCTGTCTATTTGGCAATTCGAAGGCGAAATGAGAGCAAAA 2052
Db 1809 TCCGAATGCGAGAGACGCTGTCTATTTGGCAATTCGAAGGCGAAATGAGAGCAAAA 1868
Qy 2053 AGAAGATGACAGATGATGATATCATCAGAGACGATCAAGGCTTCTGCTACGTTAG 2112
Db 1869 AGAAGATGACAGATGATGATATCATCAGAGACGATCAAGGCTTCTGCTACGTTAG 1928
Qy 2113 TCTACACAGAGAGATTCAGGCAATTAATCTCTGCAAGCGGTGGAAATGAGGTTCAATCA 2172
Db 1929 TCTACACAGAGAGATTCAGGCAATTAATCTCTGCAAGCGGTGGAAATGAGGTTCAATCA 1968
Qy 2173 AACTCTTCTTAAGGTAACCTCTGAAAGTCAATGACACAGAGCAATTTGAAAGAACTTCTCA 2232
Db 1989 AACTCTTCTTAAGGTAACCTCTGAAAGTCAATGACACAGAGCAATTTGAAAGAACTTCTCA 2048
Qy 2233 TAAAGATGATGATGAGATGAGCTCTTAAGACCAAAAGATGTCCAATATGATGACACTAG 2282
Db 2049 TAAAGATGATGATGAGATGAGCTCTTAAGACCAAAAGATGTCCAATATGATGACACTAG 2108
Qy 2293 CCAAGAGTGTGTAACAGAGACTTCATGACAGCTCATCAACACCCATCTCAACAGAT 2352
Db 2109 CCAAGAGTGTGTAACAGAGACTTCATGACAGCTCATCAACACCCATCTCAACAGAT 2168
Qy 2353 GGATGATTTCTGTGAACAAGTTTGGAAAGGACCGAAACCAACGTGGCAAAAGGCCAGG 2412
Db 2169 GGATGATTTCTGTGAACAAGTTTGGAAAGGACCGAAACCAACGTGGCAAAAGGCCAGG 2228


```

Db      1029 AGCCCTGTATGTATAGCATAGTGTAGAGAGGGGTCTCTTGGTCATATAGCCCA 1088
Qy      1273 CAGGGATGGAACCACTATCAATGGGTGCTTATCAAGGAGAGTCCCTATCCAGGCGC 1332
Db      1089 CAGGGATGGAACCACTATCAATGGGTGCTTATCAAGGAGAGTCCCTATCCAGGCGC 1148
Qy      1333 AGGAACCTGTCCAGCAAAACATTTGGTGTTCCTCTCAAAAGACCTTCTGTATGA 1392
Db      1149 AGGAACCTGTCCAGCAAAACATTTGGTGTTCCTCTCAAAAGACCTTCTGTATGA 1208
Qy      1393 TGTATTAACCTTTGCAAGAGTCAATCAGCCATGTACAAATCCAGTGTTCCTATGA 1452
Db      1209 TGTATTAACCTTTGCAAGAGTCAATCAGCCATGTACAAATCCAGTGTTCCTATGA 1268
Qy      1453 TCGCCCAATAGTATCAAAAGAGTAAATTAATTAATTAATTAATTAATTAATTAAT 1512
Db      1289 TCGCCCAATAGTATCAAAAGAGTAAATTAATTAATTAATTAATTAATTAATTAAT 1328
Qy      1513 AGTGATGCAAGAGTGAAGAGTATGATGATGATGATGATGATGATGATGATGATGAT 1572
Db      1329 AGTGATGCAAGAGTGAAGAGTATGATGATGATGATGATGATGATGATGATGATGAT 1388
Qy      1573 TCTTAAAGTATGTTCAATTCCTAAGAGACCTTGGTATGATGATGATGATGATGATGAT 1632
Db      1389 TCTTAAAGTATGTTCAATTCCTAAGAGACCTTGGTATGATGATGATGATGATGATGAT 1448
Qy      1633 AGAATGACAGTGTTCGGGAAACGAGCTGATATTCAGCAATGAGGCTTCCACATGA 1692
Db      1449 AGAATGACAGTGTTCGGGAAACGAGCTGATATTCAGCAATGAGGCTTCCACATGA 1508
Qy      1693 GCACCACTATATATATGTTTCAACGCGCTGGGTTCCAGACTCCCTTACACCGGTGA 1752
Db      1509 GCACCACTATATATATGTTTCAACGCGCTGGGTTCCAGACTCCCTTACACCGGTGA 1568
Qy      1753 TATTTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGACCTTACCTGCTTGGGA 1812
Db      1569 TATTTACGGGAAAGCGTGTGCTGAGTGTGCTGCGCCGAGACCTTACCTGCTTGGGA 1628
Qy      1813 TGGTTCGTCATGTTCTGCTATTTTCCAGTCGCAAGAGAGCAACAAAGCAAGATAT 1872
Db      1629 TGGTTCGTCATGTTCTGCTATTTTCCAGTCGCAAGAGAGCAACAAAGCAAGATAT 1688
Qy      1873 AAGAATGAGAGCCCACTGACTCACTGTTAGACTTACCAATGATATCAACATGAGCA 1932
Db      1689 AAGAATGAGAGCCCACTGACTCACTGTTAGACTTACCAATGATATCAACATGAGCA 1748
Qy      1933 CAGCCCTGAAGAGAAATCATCTATGTGTAGAGAAATGATGACACATTTTGGAAATGAG 1992
Db      1749 CAGCCCTGAAGAGAAATCATCTATGTGTAGAGAAATGATGACACATTTTGGAAATGAG 1808
Qy      1993 TCCGAAGTCGAGAGAGCGGTGTATTTGGCAATTCAGAGGCGCAATGAAAGCGGAA 2052
Db      1809 TCCGAAGTCGAGAGAGCGGTGTATTTGGCAATTCAGAGGCGCAATGAAAGCGGAA 1868
Qy      2053 AGAAGAGATCAGAGTGAATCATATCATCATGAGACAGATCAAGGCTTCTGCTACGTAG 2112
Db      1869 AGAAGAGATCAGAGTGAATCATATCATCATGAGACAGATCAAGGCTTCTGCTACGTAG 1928
Qy      2113 TCTACACAGAGAGATTCAGGCAATTCCTGCGATGCGGTGGAACATGGGTTTATACA 2172
Db      1929 TCTACACAGAGAGATTCAGGCAATTCCTGCGATGCGGTGGAACATGGGTTTATACA 1988
Qy      2173 AACTCTCTTAAAGGTAACCTCGAGAGTCAATGACACAGAGCTTTGGAAGAACTTTCTCA 2232
Db      1989 AACTCTCTTAAAGGTAACCTCGAGAGTCAATGACACAGAGCTTTGGAAGAACTTTCTCA 2048
Qy      2233 TAAAGATGATGAGAGTGGCTCTAAGACCAAGAAATGTCATATGATGATGATGATGATGAT 2292
Db      2049 TAAAGATGATGAGAGTGGCTCTAAGACCAAGAAATGTCATATGATGATGATGATGATGAT 2108
Qy      2293 CCAAGAGTCTGTGACAGAGACTTCATGACGCTCAACCAACCCCAATCTCAACAGCAT 2352

```

```

Db      2109 CCAGAGGCTGTGTACAGAGACTTCATGACAGCTCATCAACCCCAATCTCAACAGAT 2168
Qy      2253 GGATGAGTGTGTGTGAACAAGTTTGAAAGGGAACGAAACCAAGTGGGCAAGGCCAGG 2412
Db      2169 GGATGAGTGTGTGTGAACAAGTTTGAAAGGGAACGAAACCAAGTGGGCAAGGCCAGG 2228
Qy      2413 ACATACCCAGAGGAACAGTAAACAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAA 2472
Db      2229 ACATACCCAGAGGAACAGTAAACAATGGAAGCACTTACAGAAATTAAGAAAGGTAGAA 2288
Qy      2473 CAGAGAGCCCAAGATTTTGAAGAGGCAACCAAGAGTGTGAGCTGATTAACCTTGA 2532
Db      2289 CAGAGAGCCCAAGATTTTGAAGAGGCAACCAAGAGTGTGAGCTGATTAACCTTGA 2348
Qy      2533 AACCTCAACAGAGTGAAGTGTGCTGAGCAATTAAGTGAAGAAACAATGCAATTAATCAT 2592
Db      2349 AACCTCAACAGAGTGAAGTGTGCTGAGCAATTAAGTGAAGAAACAATGCAATTAATCAT 2408
Qy      2593 GAACTTTTTCATGAGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2652
Db      2409 GAACTTTTTCATGAGCATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2468
Qy      2653 CCATTTAATTAATTAATTCATGAGTAATCTTCTTAATAGCTTTTCTT 2700
Db      2469 CCAATTTAATTAATTAATTCATGAGTAATCTTCTTAATAGCTTTTCTT 2516

RESULT 6
PCT-US94-10151A-53
Sequence 53, Application PC/US9410151A
GENERAL INFORMATION:
APPLICANT: The Regents of the University of California
TITLE OF INVENTION: The Semaphorin Gene Family
NUMBER OF SEQUENCES: 66
CORRESPONDENCE ADDRESS:
ADDRESS: FLEHR HOBBACH TEST ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/10151A
FILING DATE: 13-SEP-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Osmen, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: PP-58750-PC/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299 FRT UR
INFORMATION FOR SEQ ID NO: 53:
SEQUENCE CHARACTERISTICS:
LENGTH: 2601 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 16..2331
PCT-US94-10151A-53

Query Match 92.6%; Score 2508; DB 5; Length 2601;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```


QY 193 CTGACGATGGGCTGGTAACTAGATGTGTCTTTCTGGGAGATTACTTACAGC 252
DB 9 CTGACGATGGGCTGGTAACTAGATGTGTCTTTCTGGGAGATTACTTACAGC 68
QY 253 AAGAGCAAACTATCAGATGGGAGAAACAATGTGCCAGGCTGAATTAATCTCAACAAG 312
DB 69 AAGAGCAAACTATCAGATGGGAGAAACAATGTGCCAGGCTGAATTAATCTCAACAAG 128
QY 313 AATGTGGAATCCAAATGTGATCACTTCAATGGCTGGCCAAAGCTCCAGTATCA 372
DB 129 AATGTGGAATCCAAATGTGATCACTTCAATGGCTGGCCAAAGCTCCAGTATCA 188
QY 373 TACCTTCTTTGGATGAGGAAAGAGTAGGCTGTATGTTGAGCAAAAGATCAATAT 432
DB 189 TACCTTCTTTGGATGAGGAAAGAGTAGGCTGTATGTTGAGCAAAAGATCAATAT 248
QY 433 TTCATTGCACTGGTAAATATCAAGATTTTCAAAAGATTGTGTGCCAGTATCTTAAC 492
DB 249 TTCATTGCACTGGTAAATATCAAGATTTTCAAAAGATTGTGTGCCAGTATCTTAAC 308
QY 493 CAGAAGATGATGCAAGTGGGCTGGAAAGACATCTGAAGATGTGCTAATTTGAT 552
DB 309 CAGAAGATGATGCAAGTGGGCTGGAAAGACATCTGAAGATGTGCTAATTTGAT 368
QY 553 CAAGTACTTAAGGCATATATCACTGATGATGAGCTGTGAAAGCGGGGCTTTTCA 612
DB 369 CAAGTACTTAAGGCATATATCACTGATGATGAGCTGTGAAAGCGGGGCTTTTCA 428
QY 613 TCCAAATTTGCACTTACATTTGAATTTGACATCATCTTGAGACAAATTTTAAAGTGA 672
DB 429 TCCAAATTTGCACTTACATTTGAATTTGACATCATCTTGAGACAAATTTTAAAGTGA 488
QY 673 GAACCTCACATTTTGAAGAGCGGCGTGGAGAGTCCATATGACCCCTAGCTGTGACGC 732
DB 489 GAACCTCACATTTTGAAGAGCGGCGTGGAGAGTCCATATGACCCCTAGCTGTGACGC 548
QY 733 ATCCCTTTAATAGATGAGAAATTAATCTGTGAACTGCAAGCTGATTTTATGGGCGAGA 792
DB 549 ATCCCTTTAATAGATGAGAAATTAATCTGTGAACTGCAAGCTGATTTTATGGGCGAGA 608
QY 793 CTTTGCTATCTTCCGAACCTTTGGGCAACAACCCATACAGACAGAGCATGATTC 852
DB 609 CTTTGCTATCTTCCGAACCTTTGGGCAACAACCCATACAGACAGAGCATGATTC 668
QY 853 CAGGTGGCTCAATGATCCAAAGTTCAATGAGTGGCCACTCATCTGAGAGTGAACAATCC 912
DB 669 CAGGTGGCTCAATGATCCAAAGTTCAATGAGTGGCCACTCATCTGAGAGTGAACAATCC 728
QY 913 TGAAGATGCAAAATATACCTTTTCTCCGTGAAAATGCAATAGATGAGAACTCTGG 972
DB 729 TGAAGATGCAAAATATACCTTTTCTCCGTGAAAATGCAATAGATGAGAACTCTGG 788
QY 973 AAAAGTACTCACTGATAGATAGTCAATATGCAAGAAATGACTTTGAGGGGCAACAAG 1032
DB 789 AAAAGTACTCACTGATAGATAGTCAATATGCAAGAAATGACTTTGAGGGGCAACAAG 848
QY 1033 TCTGTGAAATAATGAGCAAACTCTCAAGCTGCTGATTTGCTGAGTCCAGGCTCC 1092
DB 849 TCTGTGAAATAATGAGCAAACTCTCAAGCTGCTGATTTGCTGAGTCCAGGCTCC 908
QY 1093 AAATGCAATGCACTCATTTTGTATGAACTGAGAGATGTATCTTAATGAACTTTAAAG 1152
DB 909 AAATGCAATGCACTCATTTTGTATGAACTGAGAGATGTATCTTAATGAACTTTAAAG 968
QY 1153 TCTTAAAAATCCAGTTGTATATGAGAGTGTTAACGACTTCAGTAACATTTTCAAGGATC 1212
DB 969 TCTTAAAAATCCAGTTGTATATGAGAGTGTTAACGACTTCAGTAACATTTTCAAGGATC 1028
QY 1213 AGCCGTGTATGTATAGCATGATGATGTGAGAGGAGTCTCTTGGTCCATATGSCCA 1272
DB 1029 AGCCGTGTATGTATAGCATGATGATGTGAGAGGAGTCTCTTGGTCCATATGSCCA 1088

QY 1273 CAGGATGAGACCCCACTATCAATGGGTGCTTATCAAGAAAGATCCCTATCCAGGCC 1332
DB 1089 CAGGATGAGACCCCACTATCAATGGGTGCTTATCAAGAAAGATCCCTATCCAGGCC 1148
QY 1333 AGGAATCTGTCCAGCAAAACATTTGGTGTGTTTGTACTCTTACAAAGACCTTCTGATGA 1392
DB 1149 AGGAATCTGTCCAGCAAAACATTTGGTGTGTTTGTACTCTTACAAAGACCTTCTGATGA 1208
QY 1393 TGTATTAACCTTTGCAAGAGTCAATCCAGCCATGATCAATCAAGTGTTCATGAACA 1452
DB 1209 TGTATTAACCTTTGCAAGAGTCAATCCAGCCATGATCAATCAAGTGTTCATGAACA 1268
QY 1453 TCGCCCAATAGTATCAAAACGAGATTAATATCAATTAACAATTTGTGTAGACCG 1512
DB 1269 TCGCCCAATAGTATCAAAACGAGATTAATATCAATTAACAATTTGTGTAGACCG 1328
QY 1513 AGTGAATGCAAGATGAGACATGATGATGATTTATGGAACAGATGTTGGACCGT 1572
DB 1329 AGTGAATGCAAGATGAGACATGATGATGATTTATGGAACAGATGTTGGACCGT 1388
QY 1573 TCTTAAGATGTTCAATTCCTAAGAGACCTTGGATGATTTAAGAGAGTTCGTGGA 1632
DB 1389 TCTTAAGATGTTCAATTCCTAAGAGACCTTGGATGATTTAAGAGAGTTCGTGGA 1448
QY 1633 AGAAATGACAGTATTTTGGGAAACCGACTGTATTTACAGCAATGAGCTTTCACCTAGCA 1692
DB 1449 AGAAATGACAGTATTTTGGGAAACCGACTGTATTTACAGCAATGAGCTTTCACCTAGCA 1508
QY 1693 GCAACACTATATATTTGTTTCAAGGCTGGGGTTGCCAGTCCCTTTACACCGGTGGA 1752
DB 1509 GCAACACTATATATTTGTTTCAAGGCTGGGGTTGCCAGTCCCTTTACACCGGTGGA 1568
QY 1753 TATTTACGGAAGAGGTGTGCTGATGTTGCTGCGCCGAGACCTTACTGTGCTGGGA 1812
DB 1569 TATTTACGGAAGAGGTGTGCTGATGTTGCTGCGCCGAGACCTTACTGTGCTGGGA 1628
QY 1813 TGGTTCTGATGTTCTGCTATTTTCCACTGCAAGAGAGCAACAAGACCAAGATAT 1872
DB 1629 TGGTTCTGATGTTCTGCTATTTTCCACTGCAAGAGAGCAACAAGACCAAGATAT 1688
QY 1873 AAGAAATGGAACCCCACTGATCTACTGTTCAGACTTACCATATATATCAACATGGCCA 1932
DB 1689 AAGAAATGGAACCCCACTGATCTACTGTTCAGACTTACCATATATATCAACATGGCCA 1748
QY 1933 CAGCCCTGAAGAGAAATCATCTATGAGTGTGAGAAATGTGACATTTTGGAAATGAG 1992
DB 1749 CAGCCCTGAAGAGAAATCATCTATGAGTGTGAGAAATGTGACATTTTGGAAATGAG 1808
QY 1993 TCCGAAGTGCAGAGAGCGCTGTCTATTTGGCAATTCAGAGCGGAAATGAGAGCGAA 2052
DB 1809 TCCGAAGTGCAGAGAGCGCTGTCTATTTGGCAATTCAGAGCGGAAATGAGAGCGAA 1868
QY 2053 AGAAGATGCAAGATGATGATGATCAATCAATGAGACAGATCAAGGCTTGTGCTAGAG 2112
DB 1869 AGAAGATGCAAGATGATGATGATCAATCAATGAGACAGATCAAGGCTTGTGCTAGAG 1928
QY 2113 TCTACACAGAAAGATTCAGGCAATTAATCTGTCAGTGGGTGGAACATGGGTTCAATCA 2172
DB 1929 TCTACACAGAAAGATTCAGGCAATTAATCTGTCAGTGGGTGGAACATGGGTTCAATCA 1988
QY 2173 AACTCTTTAAGGTAAACCTGTGAAGTCAATTTGACACAGAGATTTGGAAGACTTTTCA 2232
DB 1989 AACTCTTTAAGGTAAACCTGTGAAGTCAATTTGACACAGAGATTTGGAAGACTTTTCA 2048
QY 2233 TAAAGATGATGATGAGATGGCTTTAAGACCAAGAAATGTCCAAATGACATGACACCTAG 2292
DB 2049 TAAAGATGATGATGAGATGGCTTTAAGACCAAGAAATGTCCAAATGACATGACACCTAG 2108
QY 2293 CCAGAGGTCTGTATGAGAGACTTCATGACAGCTCATCAACCAACCCATCTCAACAGAT 2352
DB 2109 CCAGAGGTCTGTATGAGAGACTTCATGACAGCTCATCAACCAACCCATCTCAACAGAT 2168
QY 2353 GATGAGTTCGTGTAACAAGTTTGGAAAGGAGCGAAACCAAGTGGGCAAGGCGAGG 2412

Db 2169 GATAGATTCTGTGAAAGATTGGAAAAAGGACCGAAACAACTCCGCAAGCCAGG 2228
Qy 2413 ACATACCCAGGAGACAGTAACAATGGAAGCACTTACAGAAATTAAGAAAGTAGAA 2472
Db 2229 ACATACCCAGGAGACAGTAACAATGGAAGCACTTACAGAAATTAAGAAAGTAGAA 2288
Qy 2473 CAGAGAGACCAAGATTGGAAGGCAACCAAGAGTCTGAGCTGATTAACCTTGA 2532
Db 2289 CAGAGAGACCAAGATTGGAAGGCAACCAAGAGTCTGAGCTGATTAACCTTGA 2348
Qy 2533 AACCTCAACAGTAGTAAGAACTTGGCTTACAGCAATTAATGGAAGAAATGCAATTAAT 2592
Db 2349 AACCTCAACAGTAGTAAGAACTTGGCTTACAGCAATTAATGGAAGAAATGCAATTAAT 2408
Qy 2593 GAACTTTTTCATGAGCATTAATGAGTGTTCATGAGTGTTCATGAGTGTTCATGAGTGTTC 2652
Db 2409 GAACTTTTTCATGAGCATTAATGAGTGTTCATGAGTGTTCATGAGTGTTCATGAGTGTTC 2468
Qy 2653 CCAATTATTAATTAATTCATGAGTGTTCATGAGTGTTCATGAGTGTTCATGAGTGTTC 2700
Db 2469 CCAATTATTAATTAATTCATGAGTGTTCATGAGTGTTCATGAGTGTTCATGAGTGTTC 2516

RESULT 7

US-08-136-922-1
Sequence 1, Application US/08136922
Patent No. 5416197

GENERAL INFORMATION:

APPLICANT: Raper, Jonathan A.
APPLICANT: Luo, Yuling
TITLE OF INVENTION: Compositions Which Regulate Neural
TITLE OF INVENTION: Regeneration and Methods of Making and Using the Same
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:

ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: No. 5416197x18
STREET: One Liberty Place
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08136,922

FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Deluca, Mark
REGISTRATION NUMBER: 33,229
REFERENCE/DOCKET NUMBER: UPN-1428
TELECOMMUNICATION INFORMATION:

TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 1481 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: both

MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 50..1480
US-08-136-922-1

Query Match 46.3%; Score 1253; DB 1; Length 1481;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1303; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 904 TGACAAATCTGAAAGATGACAAAGTATCTTTTCTCCGTAAGAAATGATAGATGAGA 963
Db 58 TGAACAATCTGAAAGATGACAAAGTATCTTTTCTCCGTAAGAAATGATAGATGAGA 117
Qy 964 ACATCTGGAAGAAAGCTCTCAAGCTTAAATGATGATGATGATGATGATGATGATGATGAT 1023
Db 118 ACATCTGGAAGAAAGCTCTCAAGCTTAAATGATGATGATGATGATGATGATGATGATGAT 1177
Qy 1024 GCAAGAGTCTGGTGAATTAATGGAACAATCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 1083
Db 178 GCAAGAGTCTGGTGAATTAATGGAACAATCTCAAGCTCTCAAGCTCTCAAGCTCTCAAG 237
Qy 1084 GCGAGTCTCAAAATGCAATGACCTCAATTTGATGAACTGAGAGTATATCTCAATGAA 1143
Db 238 GCGAGTCTCAAAATGCAATGACCTCAATTTGATGAACTGAGAGTATATCTCAATGAA 297
Qy 1144 CTTTAAAGATCTTAAATATCAAGTGTATATGAGTGTATGAGTGTATGAGTGTATGAGTGT 1203
Db 298 CTTTAAAGATCTTAAATATCAAGTGTATATGAGTGTATGAGTGTATGAGTGTATGAGTGT 357
Qy 1204 CAAGGATCAAGCGGTGTATGATATGATGATGATGATGATGATGATGATGATGATGATGATG 1263
Db 358 CAAGGATCAAGCGGTGTATGATATGATGATGATGATGATGATGATGATGATGATGATGATG 417
Qy 1264 ATATGCCCAAGGGATGAGACCAACTATCAATGGGTCTTATCAAGAAAGATGCCCTTA 1323
Db 418 ATATGCCCAAGGGATGAGACCAACTATCAATGGGTCTTATCAAGAAAGATGCCCTTA 477
Qy 1324 TCACGGCCAGGAACCTTGTCCAGCAAAACATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1383
Db 478 TCACGGCCAGGAACCTTGTCCAGCAAAACATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 537
Qy 1384 TCCGATGATGATTAATCTTTGGAAGATGATGATGATGATGATGATGATGATGATGATGATG 1443
Db 538 TCCGATGATGATTAATCTTTGGAAGATGATGATGATGATGATGATGATGATGATGATGATG 597
Qy 1444 TATGAACAATGCGCCCAATGATGATCAAAAGATGATGATGATGATGATGATGATGATGATG 1503
Db 598 TATGAACAATGCGCCCAATGATGATCAAAAGATGATGATGATGATGATGATGATGATGATG 657
Qy 1504 GATGACCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1563
Db 658 GATGACCGAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 717
Qy 1564 TGGGACCGTCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1623
Db 718 TGGGACCGTCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 777
Qy 1624 TCTGCTGGAAGAAATGACAGTTTTCGGAACGACGCTATTTTCAGCAATGAGCTTTC 1683
Db 778 TCTGCTGGAAGAAATGACAGTTTTCGGAACGACGCTATTTTCAGCAATGAGCTTTC 837
Qy 1684 CACTAAGACCAACCAATATATATGATGATGATGATGATGATGATGATGATGATGATGATG 1743
Db 838 CACTAAGACCAACCAATATATATGATGATGATGATGATGATGATGATGATGATGATGATG 897
Qy 1744 CCGGTGTGATATTTAAGGGAAGGCTGTGATGATGATGATGATGATGATGATGATGATGATG 1803
Db 898 CCGGTGTGATATTTAAGGGAAGGCTGTGATGATGATGATGATGATGATGATGATGATGATG 957
Qy 1804 TGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1863
Db 958 TGGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1017
Qy 1864 ACAAAGATATTAAGAAATGAGACCACTGACTGATGATGATGATGATGATGATGATGATGAT 1923
Db 1018 ACAAAGATATTAAGAAATGAGACCACTGACTGATGATGATGATGATGATGATGATGATGAT 1077
Qy 1924 CCAATGCGCAAGCCCTGGAAGAGAAATCATCTATGATGATGATGATGATGATGATGATGATG 1983
Db 1078 CCAATGCGCAAGCCCTGGAAGAGAAATCATCTATGATGATGATGATGATGATGATGATGATG 1137

Qy 1984 GGAATGCACTCCGAAGTCCAGAGAGCGCTGTCTATTGGCAATTCCAGAGGCGAATGA 2043
Db 1138 GGAATGCACTCCGAAGTCCAGAGAGCGCTGTCTATTGGCAATTCCAGAGGCGAATGA 1197
Qy 2044 AGAGCGAAGAGAGATCAGAGTGAATGATCATATCATCAGACAGATCAAGGCTTCT 2103
Db 1198 AGAGCGAAGAGAGATCAGAGTGAATGATCATATCATCAGACAGATCAAGGCTTCT 1257
Qy 2104 GGTACGATCTTACACAGAGAGATCAGAGATTCCTCTGCGATGCGGTGGACATG 2163
Db 1258 GGTACGATCTTACACAGAGAGATCAGAGATTCCTCTGCGATGCGGTGGACATG 1317
Qy 2164 GTTCATACAACTCTTCTTAAAGTAACTCTGGAAGTCAATGACA 2207
Db 1318 GTTCATACAACTCTTCTTAAAGTAACTCTGGAAGTCAATGACA 1361

RESULT 8

US-08-786-531B-5/c
; Sequence 5, Application US/08786531B
; Patent No. 6541197
; GENERAL INFORMATION:
; APPLICANT: Link, Charles J.
; APPLICANT: Levy, John P.
; APPLICANT: Wang, Suning
; APPLICANT: Seregina, Tatiana
; TITLE OF INVENTION: Vehicles for Stable Transfer of Green
; TITLE OF INVENTION: Fluorescent Protein Gene and Methods of Use for Same
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zarley, McKee, Thomle, Voorhees & Sease
; STREET: 801 Grand Suite 3200
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/786,531B
; FILING DATE: 21-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/010371
; FILING DATE: 22-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nebel, Heidi S.
; REGISTRATION NUMBER: 37,719
; REFERENCE/DOCKET NUMBER: hgtri
; TELEPHONE: 515-288-3667
; TELEFAX: 515-288-1338
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7160 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-786-531B-5

Query Match 1.1%; Score 31; DB 4; Length 7160;
Best Local Similarity 100.0%; Pred. No. 4.4e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGTTTAAACAAGCTT 31
Db 4179 AATCTTTATTTATCGATGTTTAAACAAGCTT 4149

RESULT 9

US-08-786-531B-6/c
; Sequence 6, Application US/08786531B
; Patent No. 6541197
; GENERAL INFORMATION:
; APPLICANT: Link, Charles J.
; APPLICANT: Levy, John P.
; APPLICANT: Wang, Suning
; APPLICANT: Seregina, Tatiana
; TITLE OF INVENTION: Vehicles for Stable Transfer of Green
; TITLE OF INVENTION: Fluorescent Protein Gene and Methods of Use for Same
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zarley, McKee, Thomle, Voorhees & Sease
; STREET: 801 Grand Suite 3200
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/786,531B
; FILING DATE: 21-JAN-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/010371
; FILING DATE: 22-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nebel, Heidi S.
; REGISTRATION NUMBER: 37,719
; REFERENCE/DOCKET NUMBER: hgtri
; TELEPHONE: 515-288-3667
; TELEFAX: 515-288-1338
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7235 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-786-531B-6

Query Match 1.1%; Score 31; DB 4; Length 7235;
Best Local Similarity 100.0%; Pred. No. 4.4e-05;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATGTTTAAACAAGCTT 31
Db 4254 AATCTTTATTTATCGATGTTTAAACAAGCTT 4224

RESULT 10
US-08-786-531B-4/c
; Sequence 4, Application US/08786531B
; Patent No. 6541197
; GENERAL INFORMATION:
; APPLICANT: Link, Charles J.
; APPLICANT: Levy, John P.
; APPLICANT: Wang, Suning
; APPLICANT: Seregina, Tatiana
; TITLE OF INVENTION: Vehicles for Stable Transfer of Green
; TITLE OF INVENTION: Fluorescent Protein Gene and Methods of Use for Same
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:

```

; ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
; STREET: 801 Grand Suite 3200
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/786,531B
; FILING DATE: 21-JAN-1997
; CLASSIFICATION: 435
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 60/010371
; FILING DATE: 22-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nebel, Heidi S.
; REGISTRATION NUMBER: 37,719
; REFERENCE/DOCKET NUMBER: hgtri
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 515-288-3667
; TELEFAX: 515-288-1338
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7352 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
;
US-08-786-531B-4
;
Query Match 0.8%; Score 22; DB 4; Length 7352;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATGTT 22
Db 4371 AATCTTTATTTATCGATGTT 4350

RESULT 11
US-08-786-531B-1/C
; Sequence 1, Application US/08786531B
; GENERAL INFORMATION:
; APPLICANT: Link, Charles J.
; APPLICANT: Levy, John P.
; APPLICANT: Wang, Suning
; APPLICANT: Seregina, Tatiana
; TITLE OF INVENTION: Vehicles for Stable Transfer of Green
; TITLE OF INVENTION: Fluorescent Protein Gene and Methods of Use for Same
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Zarley, McKee, Thomte, Voorhees & Sease
; STREET: 801 Grand Suite 3200
; CITY: Des Moines
; STATE: Iowa
; COUNTRY: United States
; ZIP: 50309
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/786,531B
; FILING DATE: 21-JAN-1997
; CLASSIFICATION: 435
```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/010371
; FILING DATE: 22-JAN-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Nebel, Heidi S.
; REGISTRATION NUMBER: 37,719
; REFERENCE/DOCKET NUMBER: hgtri
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 515-288-3667
; TELEFAX: 515-288-1338
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7353 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
;
US-08-786-531B-1
;
Query Match 0.8%; Score 22; DB 4; Length 7353;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATGTT 22
Db 4372 AATCTTTATTTATCGATGTT 4351

RESULT 12
US-09-391-741A-23
; Sequence 23, Application US/09391741A
; Patent No. 6555732
; GENERAL INFORMATION:
; APPLICANT: Duvick, Jonathan P.
; APPLICANT: Sharma, Yogesh Kumar
; TITLE OF INVENTION: Rac-Like Genes and Methods of Use
; FILE REFERENCE: 0866D
; CURRENT APPLICATION NUMBER: US/09/391,741A
; CURRENT FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: 60/111,919
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 60/100,284
; PRIOR FILING DATE: 1998-09-14
; PRIOR APPLICATION NUMBER: 09/391,741
; PRIOR FILING DATE: 1999-09-08
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 1062
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (172)...(807)
;
US-09-391-741A-23
;
Query Match 0.8%; Score 21; DB 4; Length 1062;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCGGGAG 77
Db 3 GTGACCCACGCGTCGGGAG 23

RESULT 13
US-09-391-741A-33
; Sequence 33, Application US/09391741A
; Patent No. 6555732
; GENERAL INFORMATION:
; APPLICANT: Duvick, Jonathan P.
```

APPLICANT: Sharma, Yogesh Kumar
TITLE OF INVENTION: Rac-Like Genes and Methods of Use
FILE REFERENCE: 0866D
CURRENT APPLICATION NUMBER: US/09/391,741A
CURRENT FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: 60/111,919
PRIOR FILING DATE: 1998-12-11
PRIOR APPLICATION NUMBER: 60/100,284
PRIOR FILING DATE: 1998-09-14
PRIOR APPLICATION NUMBER: 09/391,741
PRIOR FILING DATE: 1999-09-08
NUMBER OF SEQ ID NOS: 51
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 33
LENGTH: 1062
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: CDS
LOCATION: (172)..(807)
US-09-391-741A-33

Query Match 0.8%; Score 21; DB 4; Length 1062;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCGACCCACGCGTCGGGAG 77
|||||
Db 3 GTCGACCCACGCGTCGGGAG 23

RESULT 14
US-09-620-312D-33
Sequence 33, Application US/09620312D

GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyun
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yunging
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghaast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIPB
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 33
LENGTH: 1181
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (173)..(874)
US-09-620-312D-33

Query Match 0.8%; Score 21; DB 4; Length 1181;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 57 GTCGACCCACGCGTCGGGAG 77
|||||
Db 59 GTCGACCCACGCGTCGGGAG 79

RESULT 15
US-08-984-288-1
Sequence 1, Application US/08984288

PATENT No. 603872
GENERAL INFORMATION:
APPLICANT: BERGSM, DEBK
TITLE OF INVENTION: NOVEL HUMAN 11CB SPLICE V
TITLE OF INVENTION: ARIANT
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSER: RATNER & PRESTIA
STREET: P.O. BOX 980
CITY: VALLEY FORGE
STATE: PA
COUNTRY: USA
ZIP: 19482
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/984,288
FILING DATE: 03-DEC-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/032,763
FILING DATE: 11-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: PRESTIA, PAUL F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: P50599
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX: 846169
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1385 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-984-288-1

Query Match 0.8%; Score 21; DB 3; Length 1385;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCGACCCACGCGTCGGGAG 77
|||||
Db 50 GTCGACCCACGCGTCGGGAG 70

RESULT 16
US-09-620-312D-664

Sequence 664, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungting
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John, Tillinghast
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 664
LENGTH: 1572
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (330)..(704)
US-09-620-312D-664

Query Match 0.8%; Score 21; DB 4; Length 1572;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACCGCGTCGGGAG 77
Db 280 GTGACCCACCGCGTCGGGAG 300

RESULT 17
US-08-815-176-2
Sequence 2, Application US/08815176
Patent No. 5874224
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Diegidio, Tony
TITLE OF INVENTION: NOVEL EGF RECEPTOR BINDING PROTEIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/815,176
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0236 US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
LIBRARY: PENNOT03
CLONE: 1291904
US-08-815-176-2

Query Match 0.8%; Score 21; DB 2; Length 1661;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACCGCGTCGGGAG 77
Db 132 GTGACCCACCGCGTCGGGAG 152

RESULT 18
US-09-197-344-2
Sequence 2, Application US/09197344
Patent No. 6417329
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Diegidio, Tony
TITLE OF INVENTION: NOVEL EGF RECEPTOR BINDING PROTEIN
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/197,344
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/815,176
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0236 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1661 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
IMMEDIATE SOURCE:
LIBRARY: PENNOT03
CLONE: 1291904
US-09-197-344-2

Query Match 0.8%; Score 21; DB 4; Length 1661;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAG 77
|||||
Db 132 GTGACCCACGCGTCCGGAG 152

RESULT 19
US-09-324-455-1
Sequence 1, Application US/09324455
Patent No. 6326481
GENERAL INFORMATION:
APPLICANT: Yowe, David
TITLE OF INVENTION: NOVEL MOLECULES OF THE AIP-RELATED
TITLE OF INVENTION: PROTEIN FAMILY AND USES THEREOF
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09324,455
FILING DATE: 02-JUN-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/087,761
FILING DATE: 02-JUN-1998
ATTORNEY/AGENT INFORMATION:
NAME: Melkielejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/069001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2026 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-324-455-1

Query Match 0.8%; Score 21; DB 4; Length 2026;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTGACCCACGCGTCCGGAG 77
|||||
Db 11 GTGACCCACGCGTCCGGAG 31

RESULT 20
US-08-850-961-31/c
Sequence 31, Application US/08850961
Patent No. 6013517
GENERAL INFORMATION:
APPLICANT: Respass, James G.
APPLICANT: De Polo, Nicholas J.
APPLICANT: Chada, Sunil
APPLICANT: Sauter, Sybille
APPLICANT: Bodner, Mordechai
APPLICANT: Driver, David A.
TITLE OF INVENTION: CROSSLESS RETROVIRAL VECTORS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:

ADDRESSEE: Chiron Corporation, Intellectual Property - R440
STREET: P.O. Box 8097
CITY: Emeryville
STATE: California
COUNTRY: USA
ZIP: 94662-8097
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/850,961
FILING DATE: 05-MAY-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kruse, No. 6013517man J.
REGISTRATION NUMBER: 35,235
REFERENCE/DOCKET NUMBER: 930049.424C4 / 1147.005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 601-3520
TELEFAX: (510) 655-3542
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-850-961-31

Query Match 0.7%; Score 20; DB 3; Length 30;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATG 20
|||||
Db 20 AATCTTTATTTATCGATG 1

RESULT 21
US-09-132-541-5/c
Sequence 5, Application US/09132541A
Patent No. 6114113
GENERAL INFORMATION:
APPLICANT: MCLAUGHLIN-TAYLOR, Elizabeth
APPLICANT: KRUGER, Mark
APPLICANT: LUNDAK, Cheryl
APPLICANT: KILLION, Catherine
TITLE OF INVENTION: HIGH EFFICIENCY GENETIC MODIFICATION METHODS
FILE REFERENCE: 1386.002
CURRENT APPLICATION NUMBER: US/09/132,541A
CURRENT FILING DATE: 1998-08-11
EARLIER APPLICATION NUMBER: 60/055,453
EARLIER FILING DATE: 1997-08-11
NUMBER OF SEQ ID NOS: 21
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 5
LENGTH: 30
TYPE: DNA
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: This information is not available
US-09-132-541-5

Query Match 0.7%; Score 20; DB 3; Length 30;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATG 20
|||||
Db 20 AATCTTTATTTATCGATG 1

```
RESULT 22
US-09-479-776-31/c
; Sequence 31, Application US/09479776
; Patent No. 6333195
; GENERAL INFORMATION:
; APPLICANT: Respass, James G.
; De Polo, Nicholas J.
; Chada, Sunil
; Sauter, Sybille
; Bodner, Mordechai
; Driver, David A.
; TITLE OF INVENTION: CROSSLESS RETROVIRAL VECTORS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESS: CHIRON CORPORATION
; STREET: INTELLECTUAL PROPERTY-R440
; P.O. BOX 8097
; CITY: EMERYVILLE
; STATE: CA
; COUNTRY: USA
; ZIP: 94662-8097
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/479,776
; FILING DATE: 07-Jan-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: KRUSE, NORMAN J.
; REGISTRATION NUMBER: 35,235
; REFERENCE/DOCKET NUMBER: 930049.424C4
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)622-4900
; TELEFAX: (206)682-6031
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 31:
US-09-479-776-31

Query Match 0.7%; Score 20; DB 4; Length 30;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 1 AATCTTTATTTTATCGATG 20
Db 20 AATCTTTATTTTATCGATG 1

RESULT 23
US-09-244-111-5
; Sequence 5, Application US/09244111
; Patent No. 6566498
; GENERAL INFORMATION:
; APPLICANT: Ni, et al.
; TITLE OF INVENTION: Human Serine Protease and Serpin Polypeptides
; FILE REFERENCE: PF391
; CURRENT APPLICATION NUMBER: US/09/244,111
; CURRENT FILING DATE: 1999-02-04
; EARLIER APPLICATION NUMBER: 60/073,961
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 840
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (115)...(603)
US-09-244-111-5

Query Match 0.7%; Score 20; DB 4; Length 840;
Best Local Similarity 100.0%; Pred. No. 8.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 57 GTGACCCACGCGTCGGGGA 76
Db 4 GTGACCCACGCGTCGGGGA 23

RESULT 24
US-09-391-741A-9
; Sequence 9, Application US/09391741A
; Patent No. 6555732
; GENERAL INFORMATION:
; APPLICANT: Davick, Jonathan P.
; TITLE OF INVENTION: Rac-Like Genes and Methods of Use
; FILE REFERENCE: 0866D
; CURRENT APPLICATION NUMBER: US/09/391,741A
; CURRENT FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: 60/111,919
; PRIOR FILING DATE: 1998-12-11
; PRIOR APPLICATION NUMBER: 60/100,284
; PRIOR FILING DATE: 1998-09-14
; PRIOR APPLICATION NUMBER: 09/391,741
; PRIOR FILING DATE: 1999-09-08
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 1059
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (169)...(804)
US-09-391-741A-9

Query Match 0.7%; Score 20; DB 4; Length 1059;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Cy 58 TCGACCCACGCGTCGGGAG 77
Db 1 TCGACCCACGCGTCGGGAG 20

RESULT 25
US-09-620-312D-562
; Sequence 562, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aйдong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungang
; APPLICANT: Wang, Duntui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
```

TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
FILE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP23
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_fl_genes Version 1.0
SEQ ID NO: 562
LENGTH: 1168
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (148)..(615)
US-09-620-312D-562

Query Match 0.7%; Score 20; DB 4; Length 1168;
Best Local Similarity 100.0%; Pred. No. 8.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACGCGTCCGGGA 76
DB 25 GTGACCCACGCGTCCGGGA 44

RESULT 26
US-08-985-950-5
Sequence 5, Application US/08985950
Patent No. 6140076

GENERAL INFORMATION:

APPLICANT: Adema, Gosee Jan
TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes;
NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: DNAX Research Institute
STREET: 901 California Avenue

CITY: Palo Alto
STATE: California

COUNTRY: USA
ZIP: 94304-1104

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/985,950
FILING DATE: 05-DEC-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,279
FILING DATE: 21-MARCH-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/033,181
FILING DATE: 16-DEC-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/032,252
FILING DATE: 06-DEC-1996

ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.

REGISTRATION NUMBER: 34,090

REFERENCE/DOCKET NUMBER: DX0670K

TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)852-9196

TELEFAX: (650)496-1204

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1279 base pairs
TYPE: nucleic acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 155..1015
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1247
OTHER INFORMATION: /note="nucleotide 1247 designated
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 218..1015
US-08-985-950-5

Query Match 0.7%; Score 20; DB 3; Length 1279;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACGCGTCCGGGA 76
DB 20 GTGACCCACGCGTCCGGGA 39

RESULT 27
US-09-546-049-5
Sequence 5, Application US/09546049
Patent No. 6479638

GENERAL INFORMATION:

APPLICANT: Adema, Gosee Jan
Meygaard, Linde
Gorman, Daniel M.

McClanahan, Terrill K.
Zurawski, Sandra M.

Zurawski, Gerard
Lanier, Lewis L.

Phillips Jr., Joseph H.

TITLE OF INVENTION: Isolated Mammalian Monocyte Cell Genes;
Related Reagents

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: DNAX Research Institute
STREET: 901 California Avenue

CITY: Palo Alto
STATE: California

COUNTRY: USA
ZIP: 94304-1104

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/546,049
FILING DATE: 10-APR-2000

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/985,950
FILING DATE: 05-DEC-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/041,279
FILING DATE: 21-MARCH-1997

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/033,181
FILING DATE: 16-DEC-1996

ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.

REGISTRATION NUMBER: 34,090

REFERENCE/DOCKET NUMBER: DX0670K

TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)852-9196

TELEFAX: (650)496-1204

INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:

LENGTH: 1279 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 155..1015
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1247
OTHER INFORMATION: /note= "nucleotide 1247 designated
C, but may be C or T"
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 218..1015
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-546-049-5

Query Match 0.7%; Score 20; DB 4; Length 1279;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTCCAGCCACCGCTCCGGGA 76
|||||
Db 20 GTCCAGCCACCGCTCCGGGA 39

RESULT 28
US-09-336-536-1
Sequence 1, Application US/09336536
Patent No. 6406884
GENERAL INFORMATION:
APPLICANT: Leiby, K.
APPLICANT: McKay, C.
APPLICANT: Bossone, S.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-144
CURRENT APPLICATION NUMBER: US/09/336,536
CURRENT FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 1338
TYPE: DNA
ORGANISM: Homo sapiens
US-09-336-536-1

Query Match 0.7%; Score 20; DB 4; Length 1338;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTCCAGCCACCGCTCCGGGA 76
|||||
Db 1 GTCCAGCCACCGCTCCGGGA 20

RESULT 29
US-08-999-774A-11
Sequence 11, Application US/08999774A
Patent No. 6274312
GENERAL INFORMATION:
APPLICANT: Gish, Kurt C.
APPLICANT: Seghezzi, Wolfgang
APPLICANT: Shanahan, Frances
APPLICANT: Lees, Emma M.
APPLICANT: McClanahan, Terrill K.
TITLE OF INVENTION: Intracellular Regulatory Molecules;
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: DNAX Research Institute
STREET: 901 California Avenue

CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/999,774A
FILING DATE: 10-DEC-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/032,818
FILING DATE: 11-DEC-1996
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: DX0646
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 852-9196
TELEFAX: (650) 496-1200
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1503 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 25..1503
US-08-999-774A-11

Query Match 0.7%; Score 20; DB 3; Length 1503;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTCCAGCCACCGCTCCGGGA 76
|||||
Db 1 GTCCAGCCACCGCTCCGGGA 20

RESULT 30
US-09-482-273-96
Sequence 96, Application US/09482273
Patent No. 6534631
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: 71 Human Secreted Proteins
FILE REFERENCE: P2030P1
CURRENT APPLICATION NUMBER: US/09/482,273
CURRENT FILING DATE: 2000-01-13
EARLIER APPLICATION NUMBER: PCT/US99/15849
EARLIER FILING DATE: 1999-07-14
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/092,922
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/092,956
EARLIER FILING DATE: 1998-07-15
NUMBER OF SEQ ID NOS: 267
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 96
LENGTH: 1772
TYPE: DNA
ORGANISM: Homo sapiens
US-09-482-273-96

Query Match 0.7%; Score 20; DB 4; Length 1772;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 TCGACCCACGCGTCCGGAG 77
|||||
DB 1 TCGACCCACGCGTCCGGAG 20

RESULT 31

US-09-620-312D-7
Sequence 7, Application US/09620312D

Patent No. 6569662
GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 7
LENGTH: 1920
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (120)..(1394)
US-09-620-312D-7

Query Match 0.7%; Score 20; DB 4; Length 1920;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCGACCCACGCGTCCGGGA 76
|||||
DB 21 GTCGACCCACGCGTCCGGGA 40

RESULT 32

US-09-489-847-113

Sequence 113, Application US/09489847

Patent No. 6476195
GENERAL INFORMATION:

APPLICANT: Rosen et al
TITLE OF INVENTION: 98 Human Secreted Proteins
FILE REFERENCE: P2031p1
CURRENT APPLICATION NUMBER: US/09/489,847
CURRENT FILING DATE: 2000-01-24
PRIOR APPLICATION NUMBER: PCT/US99/17130
EARLIER FILING DATE: 1999-07-29
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
EARLIER APPLICATION NUMBER: 60/095,486
EARLIER FILING DATE: 1998-08-05

EARLIER APPLICATION NUMBER: 60/096,319
EARLIER FILING DATE: 1998-08-12
EARLIER APPLICATION NUMBER: 60/095,454
EARLIER FILING DATE: 1998-08-06
EARLIER APPLICATION NUMBER: 60/095,455
EARLIER FILING DATE: 1998-08-06
NUMBER OF SEQ ID NOS: 376
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 113
LENGTH: 2214
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (289)
OTHER INFORMATION: n equals a,t,c,g, or c
US-09-489-847-113

Query Match 0.7%; Score 20; DB 4; Length 2214;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCGACCCACGCGTCCGGGA 76
|||||
DB 12 GTCGACCCACGCGTCCGGGA 31

RESULT 33

US-09-620-312D-1046

Sequence 1046, Application US/09620312D

Patent No. 6569662
GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_FL_genes Version 1.0
SEQ ID NO 1046
LENGTH: 2225
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (260)..(1906)
US-09-620-312D-1046

Query Match 0.7%; Score 20; DB 4; Length 2225;
Best Local Similarity 100.0%; Pred. No. 8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGGCTCCGGGA 76
|||||
Db 35 GTGACCCACGGCTCCGGGA 54

RESULT 34

US-09-489-847-30
; Sequence 30, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 2227
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (289)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-489-847-30

Query Match 0.7%; Score 20; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 8;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGGCTCCGGGA 76
|||||
Db 12 GTGACCCACGGCTCCGGGA 31

RESULT 35

US-09-221-235-1
; Sequence 1, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,235
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-221-235-1

Query Match 0.7%; Score 20; DB 3; Length 4137;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 TCGACCCACGGCTCCGGAG 77
|||||
Db 1 TCGACCCACGGCTCCGGAG 20

RESULT 36

US-09-221-928-1
; Sequence 1, Application US/09221928
; Patent No. 6121030
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,928
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-221-928-1

Query Match 0.7%; Score 20; DB 3; Length 4137;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 TCGACCCACGGCTCCGGAG 77
|||||
Db 1 TCGACCCACGGCTCCGGAG 20

RESULT 37

US-09-221-527-1
; Sequence 1, Application US/09221527
; Patent No. 6148632
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,527
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-221-527-1

Query Match 0.7%; Score 20; DB 3; Length 4137;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 TCGACCCACGGCTCCGGAG 77

DB 1 TCGACCCACGCGTCCGGAG 20

RESULT 38

US-09-221-236-1
; Sequence 1, Application US/09221236
; Patent No. 6146841
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,236
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-221-236-1

Query Match 0.7%; Score 20; DB 3; Length 4137;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 TCGACCCACGCGTCCGGAG 77
|||||
DB 1 TCGACCCACGCGTCCGGAG 20

RESULT 39

US-09-221-416-1
; Sequence 1, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-221-416-1

Query Match 0.7%; Score 20; DB 3; Length 4137;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 TCGACCCACGCGTCCGGAG 77
|||||
DB 1 TCGACCCACGCGTCCGGAG 20

RESULT 40

US-09-221-245-1
; Sequence 1, Application US/09221245
; Patent No. 6180358
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,245
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: US 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-221-245-1

Query Match 0.7%; Score 20; DB 3; Length 4137;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 TCGACCCACGCGTCCGGAG 77
|||||
DB 1 TCGACCCACGCGTCCGGAG 20

RESULT 41

US-09-163-115-1
; Sequence 1, Application US/09163115A
; Patent No. 6183962
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/163,115A
; CURRENT FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-163-115-1

Query Match 0.7%; Score 20; DB 3; Length 4137;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 TCGACCCACGCGTCCGGAG 77
|||||
DB 1 TCGACCCACGCGTCCGGAG 20

RESULT 42

US-09-221-528-1
; Sequence 1, Application US/09221528
; Patent No. 6190874
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050

```

; CURRENT APPLICATION NUMBER: US/09/221,528
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-221-528-1

Query Match          0.7%; Score 20; DB 3; Length 4137;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 TCGACCCACGGCGTCGGGAG 77
Db 1 TCGACCCACGGCGTCGGGAG 20

RESULT 43
US-09-593-553-1
; Sequence 1, Application US/09593553
; Patent No. 6200770
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: NMI-050
; CURRENT APPLICATION NUMBER: US/09/593,553
; CURRENT FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: 09/163,115
; PRIOR FILING DATE: 1998-09-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-593-553-1

Query Match          0.7%; Score 20; DB 3; Length 4137;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 TCGACCCACGGCGTCGGGAG 77
Db 1 TCGACCCACGGCGTCGGGAG 20

RESULT 44
US-09-221-237-1
; Sequence 1, Application US/09221237
; Patent No. 6214597
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,237
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentln Ver. 2.0
```

```

; SEQ ID NO 1
; LENGTH: 4137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 'n' at positions 2872, 3597 and 3682 may be any nucleic acid
; NAME/KEY: CDS
; LOCATION: (297)..(1202)
US-09-221-237-1

Query Match          0.7%; Score 20; DB 3; Length 4137;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 58 TCGACCCACGGCGTCGGGAG 77
Db 1 TCGACCCACGGCGTCGGGAG 20

RESULT 45
US-08-654-737B-3/C
; Sequence 3, Application US/08654737B
; Patent No. 6274136
; GENERAL INFORMATION:
; APPLICANT: University of Southern California
; APPLICANT: Weiner, Leslie P.
; APPLICANT: McMillan, Minnie
; TITLE OF INVENTION: CONSTRUCTION AND USE OF GENES ENCODING
; TITLE OF INVENTION: PATHOGENIC EPITOPES FOR TREATMENT OF AUTOIMMUNE DISEASE
; FILE REFERENCE: 13761-703-00 US
; CURRENT APPLICATION NUMBER: US/08/654,737B
; CURRENT FILING DATE: 1996-05-29
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5865
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Retroviral vector derived from Moloney Murine
; OTHER INFORMATION: Leukemia Virus
US-08-654-737B-3

Query Match          0.7%; Score 20; DB 3; Length 5865;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 AATCTTTATTTATCGATG 20
Db 2740 AATCTTTATTTATCGATG 2721

RESULT 46
US-09-654-449-1/c
; Sequence 1, Application US/09654449
; Patent No. 6564061
; GENERAL INFORMATION:
; APPLICANT: Palmer, Michelle A.J.
; APPLICANT: Gee, Melissa
; APPLICANT: Tilletson, Bonnie
; APPLICANT: Chang, Xiao-jia
; TITLE OF INVENTION: Receptor Function Assay for G-Protein Coupled Receptors and Orphan
; TITLE OF INVENTION: Receptors by Reporter Enzyme Mutant Complementation
; FILE REFERENCE: 4085-226-27
; CURRENT APPLICATION NUMBER: US/09/654,449
; CURRENT FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: US 60/180,669
; PRIOR FILING DATE: 2000-02-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6700
```

TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Nucleotide sequence for pICAST ALC
NAME/KEY: CDS (1457)... (4486)
LOCATION: (1457)... (4486)
US-09-654-449-1

Query Match
Best Local Similarity 100.0%; Score 20; DB 4; Length 6700;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATGATG 20
|||||
DB 5876 AATCTTTATTTATGATG 5857

RESULT 47
US-09-654-449-3/c
Sequence 3, Application US/09654449
Patent No. 6564061
GENERAL INFORMATION:
APPLICANT: Palmer, Michelle A.J.
APPLICANT: Gee, Melissa
APPLICANT: Tillotson, Bonnie
APPLICANT: Chang, Xiao-Jia
TITLE OF INVENTION: Receptor Function Assay for G-Protein Coupled Receptors and Orphan
FILE REFERENCE: 4085-226-27
CURRENT APPLICATION NUMBER: US/09/654,449
PRIOR FILING DATE: 2000-09-01
PRIOR APPLICATION NUMBER: US 60/180,669
NUMBER OF SEQ ID NOS: 5
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 8518
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Nucleotide sequence for pICAST ALN
US-09-654-449-3

Query Match
Best Local Similarity 100.0%; Score 20; DB 4; Length 8518;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATGATG 20
|||||
DB 5877 AATCTTTATTTATGATG 5858

RESULT 48
US-08-487-283A-4/c
Sequence 4, Application US/08487283A
Patent No. 6355245
GENERAL INFORMATION:
APPLICANT: Evans, Mark J.
APPLICANT: Mattis, Louis A.
APPLICANT: Mueller, Eileen Elliott
APPLICANT: Nye, Steven H.
APPLICANT: Rollins, Scott
APPLICANT: Rocher, Russell P.
APPLICANT: Springhorn, Jeremy P.
APPLICANT: Squinto, Stephen P.
APPLICANT: Thomas, Thomas C.
APPLICANT: Wilkins, James A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT
OF INFLAMMATORY DISEASES
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seth A. Fidel
STREET: 25 Science Park (Alexion)

CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.4Mb storage
COMPUTER: Macintosh Centris 610
OPERATING SYSTEM: System 7
SOFTWARE: WordPerfect 3.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,283A
FILING DATE: June 7, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/236,208
FILING DATE: 02-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Seth A. Fidel
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-152.1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 776-1790
TELEFAX: (203) 772-3655
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 8540 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Circular
MOLECULE TYPE: Other nucleic acid
DESCRIPTION: Apex-3p Eukaryotic
DESCRIPTION: Expression Vector
US-08-487-283A-4

Query Match
Best Local Similarity 100.0%; Score 20; DB 4; Length 8540;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATGATG 20
|||||
DB 3509 AATCTTTATTTATGATG 3490

RESULT 49
PCT-US96-05611A-12/c
Sequence 12, Application PC/TUS9605611A
GENERAL INFORMATION:
APPLICANT: Mueller, John P.
APPLICANT: Leonardo, Michael J.
APPLICANT: McFarland, Henry F.
APPLICANT: Mattis, Louis A.
APPLICANT: Mueller, Eileen Elliott
APPLICANT: Nye, Steven H.
APPLICANT: Palfrey, Clara M.
APPLICANT: Squinto, Stephen P.
APPLICANT: Wilkins, James A.
TITLE OF INVENTION: Modified Myelin Protein Molecules
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESSEE: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 0.8 Mb storage
COMPUTER: Macintosh Centris 610
OPERATING SYSTEM: System 7
SOFTWARE: Microsoft Word 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/05611A
FILING DATE: 02-MAY-1995
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/431,644
FILING DATE: May 2, 1995
APPLICATION NUMBER: 08/431,648
FILING DATE: May 2, 1995
APPLICATION NUMBER: 08/482,114
FILING DATE: June 7, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: ALX-129
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 255 1400
TELEFAX: (203) 254 1101
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 8540 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Circular
MOLECULE TYPE: Other nucleic acid
DESCRIPTION: Apex-3p Eukaryotic
Expression Vector
PCT-US96-05611A-12

Query Match 0.7%; Score 20; DB 5; Length 8540;
Best Local Similarity 100.0%; Pred. No. 7.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AATCTTTATTTATCGATG 20
|||||
DB 3509 AATCTTTATTTATCGATG 3490

RESULT 50
US-09-313-294A-2526
Sequence 2526, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Lalugudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 2526
LENGTH: 256
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: Incyte ID No. 6476212 700552828H1
US-09-313-294A-2526

Query Match 0.7%; Score 19; DB 4; Length 256;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1587 CAATCCTAAGAGACTTG 1605
|||||
DB 58 CAATCCTAAGAGACTTG 76

RESULT 51
US-09-461-697-136
Sequence 136, Application US/09461697
Patent No. 6272974
GENERAL INFORMATION:
APPLICANT: COGENT NEUROSCIENCE, Inc.
APPLICANT: Lo, Donald C.

APPLICANT: Barney, Shawn
APPLICANT: Thomas, Mary Beth
APPLICANT: Portbury, Stuart D.
APPLICANT: Putnam, Kasturi
APPLICANT: Katz, Lawrence C.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
FILE REFERENCE: 10001-005-999
CURRENT APPLICATION NUMBER: US/09/461,697
CURRENT FILING DATE: 1999-12-14
NUMBER OF SEQ ID NOS: 466
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 136
LENGTH: 542
TYPE: DNA
ORGANISM: Homo sapiens
US-09-461-697-136

Query Match 0.7%; Score 19; DB 3; Length 542;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 58 TCGACCCACGCGTCGGGA 76
|||||
DB 1 TCGACCCACGCGTCGGGA 19

RESULT 52
US-09-620-312D-920
Sequence 920, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Weinman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungang
APPLICANT: Wang, Duntui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 6569662el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 08/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: PL_FL_genes Version 1.0
SEQ ID NO 920
LENGTH: 571
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (63)..(425)
US-09-620-312D-920

Query Match 0.7%; Score 19; DB 4; Length 571;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCCGG 75
 |||||
 Db 31 GTGACCCACGCGTCCGG 49

RESULT 53
 US-09-512-363-9
 ; Sequence 9, Application US/09512363
 ; Patent No. 6503184
 ; GENERAL INFORMATION:
 ; APPLICANT: Ni, Jian
 ; APPLICANT: Ruben, Steven M.
 ; TITLE OF INVENTION: Human Tumor Necrosis Factor Receptor-Like Proteins
 ; TITLE OF INVENTION: Tr11, Tr11SV1, and Tr11SV2
 ; FILE REFERENCE: PF396
 ; CURRENT APPLICATION NUMBER: US/09/512,363
 ; CURRENT FILING DATE: 2000-02-23
 ; EARLIER APPLICATION NUMBER: 60/063,212
 ; EARLIER FILING DATE: 1997-10-21
 ; EARLIER APPLICATION NUMBER: 09/176,200
 ; EARLIER FILING DATE: 1998-10-21
 ; EARLIER APPLICATION NUMBER: 60/121,648
 ; EARLIER FILING DATE: 1999-02-24
 ; EARLIER APPLICATION NUMBER: 60/134,172
 ; EARLIER FILING DATE: 1999-05-13
 ; EARLIER APPLICATION NUMBER: 60/144,076
 ; EARLIER FILING DATE: 1999-07-16
 ; NUMBER OF SEQ ID NOS: 28
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 9
 ; LENGTH: 581
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (291)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (293)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (295)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (319)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (402)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (424)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (452)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (440)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (471)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (483)

OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (495)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (516)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (559)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; US-09-512-363-9

Query Match 0.7%; Score 19; DB 4; Length 581;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCCGG 75
 |||||
 Db 3 GTGACCCACGCGTCCGG 21

RESULT 54
 US-09-176-200-9
 ; Sequence 9, Application US/09176200
 ; Patent No. 6509173
 ; GENERAL INFORMATION:
 ; APPLICANT: Ni, Jian
 ; APPLICANT: Ruben, Steven M.
 ; TITLE OF INVENTION: Human Tumor Necrosis Factor Receptor-Like Proteins
 ; TITLE OF INVENTION: Tr11, Tr11SV1, and Tr11SV2
 ; FILE REFERENCE: PF396
 ; CURRENT APPLICATION NUMBER: US/09/176,200
 ; CURRENT FILING DATE: 1998-10-21
 ; EARLIER APPLICATION NUMBER: 60/063,212
 ; EARLIER FILING DATE: 1997-10-21
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 9
 ; LENGTH: 581
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (291)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (293)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (295)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (319)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (402)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (424)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (440)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (471)
 ; OTHER INFORMATION: n equals a, t, g or c
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (483)

NAME/KEY: misc_feature
LOCATION: (452)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (440)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (471)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (483)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (516)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (545)
OTHER INFORMATION: n equals a, t, g or c
FEATURE:
NAME/KEY: misc_feature
LOCATION: (559)
OTHER INFORMATION: n equals a, t, g or c
US-09-176-200-9

Query Match 0.7%; Score 19; DB 4; Length 581;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 57 GTGACCCAGCGCTCGGG 75
Db 3 GTGACCCAGCGCTCGGG 21

RESULT 55
US-08-801-972-2
Sequence 2, Application US/08801972
Patent No. 5831018
GENERAL INFORMATION:
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: A NOVEL HUMAN CYTOCHROME
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/801,972
FILING DATE: Filed Herewith
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0216 US

TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 648 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: 2264814
US-08-801-972-2

Query Match 0.7%; Score 19; DB 2; Length 648;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1615 AGAAGGTTCTGCTGGA 1633
Db 243 AGAAGGTTCTGCTGGA 261

RESULT 56
US-09-178-881-2
Sequence 2, Application US/09178881
Patent No. 6083704
GENERAL INFORMATION:
APPLICANT: Goli, Surya K.
TITLE OF INVENTION: A NOVEL HUMAN CYTOCHROME B5
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/178,881
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/801,972
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0216 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 648 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: Consensus
CLONE: 2264814
US-09-178-881-2

Query Match 0.7%; Score 19; DB 3; Length 648;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1615 AGAAGGTTCTGCTGGA 1633
|||||

DB 243 AGAAGAGTTCTGCTGAA 261

RESULT: 57

US-09-205-258-63

Sequence 63, Application US/09205258

Patent No. 6525174

GENERAL INFORMATION:

APPLICANT: Young et al.

TITLE OF INVENTION: 207 Human Secreted Proteins

FILE REFERENCE: P2007P1

CURRENT APPLICATION NUMBER: US/09/205,258

EARLIER FILING DATE: 1998-12-04

EARLIER APPLICATION NUMBER: PCT/US98/11422

EARLIER FILING DATE: 1998-06-04

EARLIER APPLICATION NUMBER: 60/048,885

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,375

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,881

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,880

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,896

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,020

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,876

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,895

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,884

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,894

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,971

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,964

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,882

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,899

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,893

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,900

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,901

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,892

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,915

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,019

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,970

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,972

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,916

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,373

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,875

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,374

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,917

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,949

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,974

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,974

EARLIER APPLICATION NUMBER: 60/048,883

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,897

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,898

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,962

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,963

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,877

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,878

EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/070,923

EARLIER FILING DATE: 1997-12-18

EARLIER APPLICATION NUMBER: 60/092,921

EARLIER FILING DATE: 1998-07-15

EARLIER APPLICATION NUMBER: 60/094,657

EARLIER FILING DATE: 1998-07-30

NUMBER OF SEQ ID NOS: 1227

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 63

LENGTH: 780

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (12)

OTHER INFORMATION: n equals a,t,g, or c

FEATURE:

NAME/KEY: SITE

LOCATION: (4)

OTHER INFORMATION: n equals a,t,g, or c

FEATURE:

NAME/KEY: SITE

LOCATION: (12)

OTHER INFORMATION: n equals a,t,g, or c

FEATURE:

NAME/KEY: SITE

LOCATION: (738)

OTHER INFORMATION: n equals a,t,g, or c

FEATURE:

NAME/KEY: SITE

LOCATION: (776)

OTHER INFORMATION: n equals a,t,g, or c

US-09-205-258-63

Query Match

Best Local Similarity 100.0%; Pred. No. 25;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCCGCCACCGCTCCGGG 75

DB 28 GTCCGCCACCGCTCCGGG 46

US-09-904-615-25

Sequence 25, Application US/09904615

Patent No. 6566325

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: 49 Human Secreted Proteins

FILE REFERENCE: P2032P1

CURRENT APPLICATION NUMBER: US/09/904,615

CURRENT FILING DATE: 2001-07-16

PRIOR APPLICATION NUMBER: 09/511,554

PRIOR FILING DATE: 2000-02-23

PRIOR APPLICATION NUMBER: 60/097,917

PRIOR FILING DATE: 1998-08-25

PRIOR APPLICATION NUMBER: 60/098,634

PRIOR FILING DATE: 1998-08-31

PRIOR APPLICATION NUMBER: 60/098,634

PRIOR FILING DATE: 1998-08-31

```
NUMBER OF SEQ ID NOS: 170
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 25
LENGTH: 831
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (5)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (10)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (11)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (15)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (27)
OTHER INFORMATION: n equals a,t,g, or c
US-09-904-615-25
```

```
Query Match 0.7%; Score 19; DB 4; Length 831;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 57 GTGACCCACGCGTCGCGG 75
Db 67 GTGACCCACGCGTCGCGG 85
```

```
RESULT 59
US-09-461-697-1
Sequence 1, Application US/09461697
Patent No. 6277974
GENERAL INFORMATION:
APPLICANT: COGENT NEUROSCIENCE, Inc.
APPLICANT: Lo, Donald C.
APPLICANT: Barney, Shawn
APPLICANT: Thomas, Mary Beth
APPLICANT: Portbury, Stuart D.
APPLICANT: Putnam, Kassturi
APPLICANT: Katz, Lawrence C.
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
FILE REFERENCE: 10001-005-999
CURRENT FILING DATE: 1999-12-14
NUMBER OF SEQ ID NOS: 466
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 852
TYPE: DNA
ORGANISM: Homo sapiens
US-09-461-697-1
```

```
Query Match 0.7%; Score 19; DB 3; Length 852;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 58 TCGACCCACGCGTCGCGGA 76
Db 1 TCGACCCACGCGTCGCGGA 19
```

```
RESULT 60
US-09-257-179-20
Sequence 20, Application US/09257179
Patent No. 6410709
GENERAL INFORMATION:
```

```
APPLICANT: Ruben et al.
TITLE OF INVENTION: 29 Human Secreted Proteins
FILE REFERENCE: P201SP1
CURRENT APPLICATION NUMBER: US/09/257,179
CURRENT FILING DATE: 1999-02-25
EARLIER APPLICATION NUMBER: PCT/US98/17709
EARLIER FILING DATE: 1998-08-27
EARLIER APPLICATION NUMBER: 60/056,270
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,271
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,247
EARLIER FILING DATE: 1997-08-29
EARLIER APPLICATION NUMBER: 60/056,073
EARLIER FILING DATE: 1997-08-29
NUMBER OF SEQ ID NOS: 128
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 20
LENGTH: 932
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (43)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (458)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (465)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (494)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (501)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (515)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (582)
OTHER INFORMATION: n equals a,t,g, or c
US-09-257-179-20
```

```
Query Match 0.7%; Score 19; DB 4; Length 932;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 57 GTGACCCACGCGTCGCGG 75
Db 72 GTGACCCACGCGTCGCGG 90
```

```
RESULT 61
US-09-512-363-3
Sequence 3, Application US/09512363
Patent No. 6503184
GENERAL INFORMATION:
APPLICANT: NI, Jlan
APPLICANT: Ruben, Steven M.
TITLE OF INVENTION: Human Tumor Necrosis Factor Receptor-Like Proteins
FILE REFERENCE: PR396
CURRENT APPLICATION NUMBER: US/09/512,363
CURRENT FILING DATE: 2000-02-23
EARLIER APPLICATION NUMBER: 60/063,212
```

```

; EARLIER FILING DATE: 1997-10-21
; EARLIER APPLICATION NUMBER: 09/176,200
; EARLIER FILING DATE: 1998-10-21
; EARLIER APPLICATION NUMBER: 60/121,648
; EARLIER FILING DATE: 1999-02-24
; EARLIER APPLICATION NUMBER: 60/134,172
; EARLIER FILING DATE: 1999-05-13
; EARLIER APPLICATION NUMBER: 60/144,076
; EARLIER FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 3
; LENGTH: 1007
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (121)..(843)
US-09-512-363-3
```

```

Query Match
Best Local Similarity 100.0%; Score 19; DB 4; Length 1007;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 57 GTGACCCACGCGTCGGG 75
Db 1 GTGACCCACGCGTCGGG 19
```

```

RESULT 62
US-09-176-200-3
; Sequence 3, Application US/09176200
; Patent No. 6509173
; GENERAL INFORMATION:
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: Human Tumor Necrosis Factor Receptor-Like Proteins
; FILE REFERENCE: PFI396
; CURRENT FILING DATE: 1998-10-21
; EARLIER FILING DATE: 1998-10-21
; EARLIER APPLICATION NUMBER: 60/063,212
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 3
; LENGTH: 1007
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (121)..(843)
US-09-176-200-3
```

```

Query Match
Best Local Similarity 100.0%; Score 19; DB 4; Length 1007;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 57 GTGACCCACGCGTCGGG 75
Db 1 GTGACCCACGCGTCGGG 19
```

```

RESULT 63
US-09-620-312D-866
; Sequence 866, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
```

```

; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
```

```

; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aiyong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yunqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillmanhast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 656962e1 Nucleic Acids and
; FILE REFERENCE: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT FILING DATE: 2000-07-19
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_fl_genes Version 1.0
; SEQ ID NO: 866
; LENGTH: 1024
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (519)..(980)
US-09-620-312D-866
```

```

Query Match
Best Local Similarity 100.0%; Score 19; DB 4; Length 1024;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 57 GTGACCCACGCGTCGGG 75
Db 14 GTGACCCACGCGTCGGG 32
```

```

RESULT 64
US-09-489-847-112
; Sequence 112, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT FILING DATE: 2000-01-24
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 112
; LENGTH: 1037
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (936)
; OTHER INFORMATION: n equals a,t,g, or c
```

```
FEATURE:
NAME/KEY: SITE
LOCATION: (946)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (951)
OTHER INFORMATION: n equals a,t,g, or c
US-09-489-847-112
```

```
Query Match          0.7%; Score 19; DB 4; Length 1037;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 57 GTCCAGCCACGCGTCGGG 75
Db 19 GTCCAGCCACGCGTCGGG 37
```

```
RESULT 65
US-09-372-422A-41
Sequence 41, Application US/09372422A
Patent No. 6313375
GENERAL INFORMATION:
APPLICANT: Rudolf Jung
TITLE OF INVENTION: Maize Aquaporins and Uses Thereof
FILE REFERENCE: 0919
CURRENT APPLICATION NUMBER: US/09/372,422A
CURRENT FILING DATE: 1999-08-11
PRIOR APPLICATION NUMBER: US 60/098,692
PRIOR FILING DATE: 1998-08-31
NUMBER OF SEQ ID NOS: 49
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 41
LENGTH: 1116
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: CDS
LOCATION: (2)...(719)
US-09-372-422A-41
```

```
Query Match          0.7%; Score 19; DB 4; Length 1116;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 57 GTCCAGCCACGCGTCGGG 75
Db 6 GTCCAGCCACGCGTCGGG 24
```

```
RESULT 66
US-09-620-312D-162
Sequence 162, Application US/09620312D
Patent No. 6569662
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungqing
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
```

```
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 656962e1 Nucleic Acids and
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 162
LENGTH: 1221
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (263)..(862)
US-09-620-312D-162
```

```
Query Match          0.7%; Score 19; DB 4; Length 1221;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 57 GTCCAGCCACGCGTCGGG 75
Db 19 GTCCAGCCACGCGTCGGG 37
```

```
RESULT 67
US-09-620-312D-972
Sequence 972, Application US/09620312D
Patent No. 6569662
```

```
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Ren, Feiyan
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungqing
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 656962e1 Nucleic Acids and
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 972
LENGTH: 1281
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1172)..(1074)
US-09-620-312D-972
```

```
Query Match          0.7%; Score 19; DB 4; Length 1281;
```

Best Local Similarity 100.0%; Pred. No. 24;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACGGGTCGGG 75
|||||
DB 14 GTGACCCACGGGTCGGG 32

RESULT 68

US-09-620-312D-196

/ Sequence 196, Application US/09620312D
/ Patent No. 6569662
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Chen, Rui-hong
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Wehrman, Tom
/ APPLICANT: Xue, Aidong J.
/ APPLICANT: Yang, Yonghong
/ APPLICANT: Wang, Jian-Rui
/ APPLICANT: Zhou, Ping
/ APPLICANT: Ma, Yungqing
/ APPLICANT: Wang, Dunrui
/ APPLICANT: Wang, Zhiwei
/ APPLICANT: John Tillinghast
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
/ FILE REFERENCE: Polypeptides
/ CURRENT APPLICATION NUMBER: US/09/620,312D
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1105
/ SOFTWARE: pt_fl_genes Version 1.0
/ SEQ ID NO 196
/ LENGTH: 1332
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (171)..(920)
US-09-620-312D-196

Query Match

0.7%; Score 19; DB 4; Length 1332;

Best Local Similarity 100.0%; Pred. No. 24;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACGGGTCGGG 75
|||||
DB 40 GTGACCCACGGGTCGGG 58

RESULT 69

US-09-620-312D-308

/ Sequence 308, Application US/09620312D
/ Patent No. 6569662
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Chen, Rui-hong
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Wehrman, Tom
/ APPLICANT: Xue, Aidong J.

/ APPLICANT: Yang, Yonghong
/ APPLICANT: Wang, Jian-Rui
/ APPLICANT: Zhou, Ping
/ APPLICANT: Ma, Yungqing
/ APPLICANT: Wang, Dunrui
/ APPLICANT: Wang, Zhiwei
/ APPLICANT: John Tillinghast
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
/ FILE REFERENCE: Polypeptides
/ CURRENT APPLICATION NUMBER: US/09/620,312D
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1105
/ SOFTWARE: pt_fl_genes Version 1.0
/ SEQ ID NO 308
/ LENGTH: 1354
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (226)..(1251)
US-09-620-312D-308

Query Match

0.7%; Score 19; DB 4; Length 1354;

Best Local Similarity 100.0%; Pred. No. 24;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCACGGGTCGGG 75
|||||
DB 129 GTGACCCACGGGTCGGG 147

RESULT 70

US-09-620-312D-871

/ Sequence 871, Application US/09620312D
/ Patent No. 6569662
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyan
/ APPLICANT: Chen, Rui-hong
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Wehrman, Tom
/ APPLICANT: Xue, Aidong J.
/ APPLICANT: Yang, Yonghong
/ APPLICANT: Wang, Jian-Rui
/ APPLICANT: Zhou, Ping
/ APPLICANT: Ma, Yungqing
/ APPLICANT: Wang, Dunrui
/ APPLICANT: Wang, Zhiwei
/ APPLICANT: John Tillinghast
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
/ FILE REFERENCE: Polypeptides
/ CURRENT APPLICATION NUMBER: US/09/620,312D
/ PRIOR FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1105
/ SOFTWARE: pt_fl_genes Version 1.0
/ SEQ ID NO 871
/ LENGTH: 1368
/ TYPE: DNA

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (103)..(447)
US-09-620-312D-871

Query Match 0.7%; Score 19; DB 4; Length 1368;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCAGCGGTCGGG 75
Db 48 GTGACCCAGCGGTCGGG 66

RESULT 71
US-09-205-258-124
Sequence 124, Application US/09205258
Patent No. 6525174
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P2007P1
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: US/09/205,258
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: PCT/US98/11422
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 124
LENGTH: 1390
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (498)
OTHER INFORMATION: n equals a,t,g, or c
FEATURE:
NAME/KEY: SITE
LOCATION: (499)
OTHER INFORMATION: n equals a,t,g, or c
US-09-205-258-124

Query Match 0.7%; Score 19; DB 4; Length 1390;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 57 GTGACCCAGCGGTCGGG 75
Db 68 GTGACCCAGCGGTCGGG 86

RESULT 72
US-08-964-127-5
Sequence 5, Application US/08964127
Patent No. 6277565
GENERAL INFORMATION:
APPLICANT: Grandearl, Andrew David John
TITLE OF INVENTION: NOVEL GENES ENCODING TRANSPORTER-LIKE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/964,127
FILING DATE: 06-NOV-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Crews, Ph.D., L. Lee
REGISTRATION NUMBER: P-43,567
REFERENCE/DOCKET NUMBER: 07334/038001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1411 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...966
US-09-964-127-5

Query Match 0.7%; Score 19; DB 3; Length 1411;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCCGGG 75
|||||
Db 1 GTGACCCACGCGTCCGGG 19

RESULT 73
US-09-496-692-5
Sequence 5, Application US/09496692
Patent No. 6313271
GENERAL INFORMATION:
APPLICANT: Grandearl, Andrew David John
TITLE OF INVENTION: NOVEL GENES ENCODING TRANSPORTER-LIKE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/496,692
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/964,127
FILING DATE: 06-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: Crews, Ph.D., L. Lee
REGISTRATION NUMBER: P-43,567
REFERENCE/DOCKET NUMBER: 07334/038001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070

TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1411 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...966
US-09-496-692-5

Query Match 0.7%; Score 19; DB 4; Length 1411;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 57 GTGACCCACGCGTCCGGG 75
|||||
Db 1 GTGACCCACGCGTCCGGG 19

RESULT 74
US-10-000-273-5
Sequence 5, Application US/10000273
Patent No. 6573057
GENERAL INFORMATION:
APPLICANT: Grandearl, Andrew David John
TITLE OF INVENTION: NOVEL GENES ENCODING TRANSPORTER-LIKE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSER: Fish & Richardson P.C.
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/000,273
FILING DATE: 02-NOV-2001
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/964,127
FILING DATE: 06-NOV-1997
ATTORNEY/AGENT INFORMATION:
NAME: Crews, Ph.D., L. Lee
REGISTRATION NUMBER: P-43,567
REFERENCE/DOCKET NUMBER: 07334/038001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1411 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...966
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-000-273-5

Query Match 0.7%; Score 19; DB 4; Length 1411;
Best Local Similarity 100.0%; Pred. No. 24;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCGACCCACGCGTCCGG 75
Db 1 GTCGACCCACGCGTCCGG 19

RESULT 75

US-09-118-442-5
; Sequence 5, Application US/09118442B
; Patent No. 6197561
; GENERAL INFORMATION:
; APPLICANT: Martino-Catt, Susan J.
; APPLICANT: Wang, Hongyu
; APPLICANT: Beach, Larry R.
; APPLICANT: Wang, Xun
; TITLE OF INVENTION: Genes Controlling Phytochrome Metabolism in
; FILE REFERENCE: 0706
; CURRENT APPLICATION NUMBER: US/09/118,442B
; EARLIER FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/055,446
; EARLIER FILING DATE: 1997-08-11
; EARLIER APPLICATION NUMBER: 60/055,526
; EARLIER FILING DATE: 1997-08-08
; EARLIER APPLICATION NUMBER: 60/053,944
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 1428
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (118)...(1176)
US-09-118-442-5

Query Match

Best Local Similarity 0.7%; Score 19; DB 3; Length 1428;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 57 GTCGACCCACGCGTCCGG 75
Db 6 GTCGACCCACGCGTCCGG 24

Search completed: July 31, 2003, 23:55:46
Job time : 188 secs

THIS PAGE BLANK (USPTO)